# DOCUMENT RESUME

ED 127 551 CS 002 776

TITLE Artificial Intelligence and Language Comprehension.

INSTITUTION National Inst. of Education (DHEW), Washington, D.C. Basic Skills Group. Learning Div.

PUB DATE Feb 76

CONTRACT 400-75-0012; NIE-P-75-0026; NIE-P-75-0027

NOTE 123p.

EDRS PRICE MF-$0.83 HC-$6.01 Plus Postage.

DESCRIPTORS *Artificial Intelligence; *Cognitive Ability; *Comprehension; Intelligence; *Language; Language Research; *Learning Processes; Linguistics; Reading Instruction; Structural Grammar; *Thought Processes

ABSTRACT

The three papers in this volume concerning artificial intelligence and language comprehension were commissioned by the National Institute of Education to further the understanding of the cognitive processes that enable people to comprehend what they read. The first paper, "Artificial Intelligence and Language Comprehension," by Terry Winograd, outlines the potential impact of artificial intelligence on our understanding of human language comprehension. It is directed to people outside the field, concentrating on the ways in which the research differs in method and content from the work being done from other viewpoints and laying out the assumptions that have led many researchers to apply the "computational metaphor" to human language use. "Artificial Intelligence, Language, and the Study of Knowledge," by Ira Goldstein, Seymour Papert, and Marvin Minsky, develops the view that intelligence is based on the ability to use large amounts of diverse kinds of knowledge in procedural ways, rather than on the possession of a few general and uniform principles. Implications for education are discussed. The third paper is "Natural Language Understanding Systems Within the AI Paradigm: A Survey and Some Comparisons," by Yorick Wilks. (Authors/MKM)

***********************************************************************
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
***********************************************************************

ED127551

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

**AI**

# artificial intelligence and language comprehension

**nie**

u.s. department
of health,
education and welfare

national institute
of education

CS 002 776

# ARTIFICIAL INTELLIGENCE AND LANGUAGE COMPREHENSION

NATIONAL INSTITUTE OF EDUCATION
WASHINGTON, D.C.
FEBRUARY, 1976

## FOREWORD

In the summer of 1973 a group of distinguished scientists convened at Hyannis, Massachusetts to design a program of research on reading and the related language processes. The report of the Hyannis meeting[1] provided impetus for NIE's decision to establish reading as a major priority.

During 1973 and 1974 NIE organized a series of activities designed to plan in greater detail some of the research initiatives suggested by the Hyannis group. They placed emphasis on finding ways to identify research that could be helpful in explaining the cognitive processes that enable people to comprehend what they read. In particular, it seemed that understanding of reading comprehension might be improved by strengthening the already-growing collaboration between cognitive psychologists and persons in the field of artificial intelligence (AI).

Accordingly, members of the Institute staff decided to commission a set of papers that would present artificial intelligence work from a variety of perspectives and to identify as many AI researchers as we could. The three papers that follow and a directory of AI research personnel[2] are the results of that decision.

The people who contributed to this volume--Terry Winograd, Ira Goldstein, Seymour Papert, Marvin Minsky, and Yorick Wilks--deserve special thanks. They have shared with us not only the nature of their enterprise but also their enthusiasm for it.

Thanks go also to Professor Herbert A. Simon of the Carnegie-Mellon University for his efforts to help us establish contact with the community of AI scientists.

Susan Duffy's thoughtful comments on the various drafts are greatly appreciated.

Monte Penney
Learning Division, Basic Skills Group
National Institute of Education

---

[1] Miller, George A. (ed.), Linguistic Communication: Perspectives for Research, Newark, Delaware: International Reading Association, 1974, 45 pp.

[2] Walker, Donald E., Artificial Intelligence and Language Processing: A Directory of Research Personnel, Menlo Park, California: Stanford Research Institute, 1975, 81 pp.

## TABLE OF CONTENTS

Foreword i

ARTIFICIAL INTELLIGENCE AND LANGUAGE COMPREHENSION 1
Terry Winograd, Department of Computer Science and Department of Linguistics, Stanford University

ARTIFICIAL INTELLIGENCE, LANGUAGE, AND THE STUDY OF KNOWLEDGE 27
Ira Goldstein, Seymour Papert, Marvin Minsky, Artificial Intelligence Laboratory, Massachusetts Institute of Technology

NATURAL LANGUAGE UNDERSTANDING SYSTEMS WITHIN THE AI PARADIGM: A SURVEY AND SOME COMPARISONS 79
Yorick Wilks, Department of Artificial Intelligence, University of Edinburgh

# ARTIFICIAL INTELLIGENCE AND LANGUAGE COMPREHENSION

Terry Winograd

Department of Computer Science
Department of Linguistics
Stanford University

January 1975

This paper outlines the potential impact of artificial intelligence on our understanding of human language comprehension. It is directed to people outside the field, concentrating on the ways in which the research differs in method and content from the work being done from other viewpoints, laying out the assumptions that have led many researchers into applying the "computational metaphor" to human language use. It includes a discussion of potential support patterns the NIE could adopt for further work in the area, indicating the fundamental areas of concentration and strategies for promoting good research.

The research and writing of this paper was sponsored under NIE Order No. NIE-P-75-0027. The views and conclusions contained here are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of the National Institute of Education or the United States Government.

## TABLE OF CONTENTS

ARTIFICIAL INTELLIGENCE AND LANGUAGE COMPREHENSION 3

I. The Problem of Understanding Understanding 4

An Example of Human Language Comprehension
An Example of Computer Language Comprehension

II. Artificial Intelligence and Cognitive Theory 7

Advantages of the AI Approach
Limitations of the AI Approach
The Effect of Artificial Intelligence on Education

III. A Strategy for AI Research 16

Where Things Stand
Specific Research Support Areas
Areas Not Mentioned for Support

References 22

# ARTIFICIAL INTELLIGENCE AND LANGUAGE COMPREHENSION

Language is the basis for most of our educational system. Through listening, talking, and reading, a student gains the understanding and knowledge of whatever field he follows. Yet we still do not really understand what happens when a person reads or hears the simplest sentence. Psychologists, linguists, and philosophers can point out many fascinating aspects of how language is used, but there is no theory presenting a comprehensive picture of what processes are going on in the mind of the language user, or how these processes relate to other aspects of thought.

In the past few years, a new field has been created which tries to combine the ideas of psychology and computer science. It has been named "artificial intelligence" and, as that name implies, one of its major goals is programing computers to do intelligent things--to converse with human beings, play chess, recognize visual objects, solve abstract problems, and use commonsense. In the process of programing a computer to behave intelligently, we come to examine human abilities from a new perspective. Computer scientists have developed a rich set of concepts about information--what it is and how it can be stored, processed, and used. Combined with the notions of computation, these concepts form a base for building comprehensive theories of what goes on in the process of language understanding. We can build complex theories, emphasizing the integration of all the facets of thought and language, and we can test them in the form of working programs on the computer.

In this paper I will discuss the current state and future promise of artificial intelligence as a method for increasing our understanding of human language processes. Section I lays out some of the problems, discussing the kind of complexities involved in human language understanding, and indicating the current level of computer programs for carrying on a natural language dialog. Section II describes the fundamental paradigm AI provides for viewing cognitive processes. It discusses the advantages of this viewpoint, some of the criticisms of current research, and the ways in which AI research relates to work in fields such as linguistics, experimental psychology, cognitive psychology, and the theory of logic. All of these areas have made contributions to the study of language, and much of the understanding to be gained from them is complementary to what we can expect from AI research. The third section outlines a strategy for promoting research in artificial intelligence relevant to the goals of the NIE essential skills program. It briefly outlines the status of current work, and analyzes the specific needs for support and direction.

Much of the work I will describe has been called "cognitive simulation" rather than "artificial intelligence." This difference in emphasis fits closely with the viewpoint I am proposing. The gain from developing AI is not primarily in the usefulness of the programs we create, but in the set of concepts we develop, and the ways in which we can apply them to understanding human intelligence. However, in this paper, I will use the term "artificial intelligence" to cover all of the research, inasmuch as the term has been established through common usage and appears in the names of the major books, journals, projects, courses, and laboratories in the field. The reader should be careful to look past the "artificial" and see the research as a basic study of intelligent processes in whatever physical manifestation they appear.

The paper is one of several prepared for a study by the National Institute of Education to assess the feasibility of supporting research in artificial intelligence, related to the goals of the essential skills program. It is not a presentation of official policy, but represents the viewpoint and suggestions of one researcher in the field. The basic goal of the paper is to explain the motivations for studying language from an artificial intelligence viewpoint, and to suggest what we might hope to achieve. It does not attempt to survey or criticize current work in detail, but is intended as a companion to other papers being prepared with that end in mind.

## I. The Problem of Understanding Understanding

Language comprehension is one of those activities every normal person does without conscious effort. Like breathing or walking, we do it all the time without thinking about how we are doing it and only think about it explicitly when communication breaks down. In fact, consciously it does not feel that we are going through a process at all--comprehension just happens all at once. Yet, from the point of view of information processing, comprehension is an astounding feat, requiring a tremendous background of knowledge and skills. A person must know not only the rules of a language, but the subtleties of the way it conveys meaning, and facts about the subject of conversation as well.

It is a formidable challenge to tackle the problem of language comprehension in its entirety. In many ways, we are in the position of the proverbial blind men, facing a rather unruly elephant. Detailed experimental data are available for a number of specialized facets, but the overall organization of our thinking processes is not observable from a single viewpoint. In a strict empirical sense, there is not much to observe. We can watch people's eyes scanning a page or verify that they are hearing a set of sounds. We can test the final result of understanding by asking them to repeat or paraphrase the message, or to answer questions about it. But what goes on between these two endpoints is a complex process which has been impervious to the usual scientific method. Researchers in education and psychology have concentrated on fine measurements of those aspects of behavior which could be directly observed, while philosophers have speculated about the internal process in ways which have been suggestive, but haven't led to testable detailed theories of what is actually going on.

The problem isn't just that the inner workings are unobservable. More important is the way in which they intertwine issues of language form, meaning, and knowledge. We cannot successfully analyze language comprehension as the "sum" of a syntactic (grammar) process, a meaning analysis, and a reasoning system which can be studied independently. In reading the simplest sentence, a person combines all of these processes, and without understanding how they interact, we cannot understand how any one of them works.

## An Example of Human Language Comprehension

To demonstrate the magnitude of the problem, let us look at a simplification of a "reading comprehension test." An elementary school student might be asked to read a text like the following paragraph, then answer simple questions to demonstrate understanding.

> Tommy had just been given a new set of blocks. He was opening the box when Jimmy came in.
>
> Questions
>
> Who was opening the box?
> What was in the box?

Although a normal student could answer these questions without difficulty, they require a surprising amount of knowledge. The questions cannot be answered by simply repeating sentences from the passage; in that case, the first question would be answered "He was opening the box." The reader has to recognize that in this context "he" means "Tommy." Thus he must know that the word "he" always refers to a person, that "Tommy" is a person's name, and that blocks are not people. All of this seems nonsensically obvious, but our experience in programing computers to use language has taught us the necessity of carefully understanding what goes into an "obvious" thought. In order to get a computer to answer this question, we must build complex models of memory and deduction making use of these facts.

The second question "What was in the box?" calls for more specialized knowledge. Nothing in the text directly says what box is being opened, or what its contents are. A person reading the text immediately assumes that the box contains the set of blocks. The reader knows that new items often come in boxes, and that opening such a box is the usual thing to do. Most importantly, one assumes one is receiving a connected message. There is no logical reason why the box must be connected with the blocks, but if it were not, a writer would not mention it without further introduction. The sequences "Tommy had just been given a new job., He was opening the box . . .," or "Tommy had just been given a new set of blocks. He was opening the bottle . . ." are confusing, since there is no obvious connection between the sentences.

In comprehending language, a person is always combining commonsense knowledge with knowledge of word meanings and the structure of sentences in language. When we communicate to another person, we need not fill in all the connections between things. We can safely assume that our listener will use intelligence and a store of knowledge to put together the message we intend. An utterance is like a rough sketch, allowing the hearer to reconstruct the meaning. This interaction between language and reasoning is central to the way understanding is structured. If we want to apply a scientific understanding of language to education, we must have theories that deal directly with these thinking processes.

## An Example of Computer Language Comprehension

One way of assessing how much we know about language comprehension is to ask how well we can duplicate it in formal systems. Much of this paper is devoted to explaining the advantages of such an approach, and in order to give readers a feeling for what has already been done, I include a segment from a dialog with a computer program designed to interact with a human in natural language. The program is a simulated robot, which manipulates a set of toy objects on a table top. It conducts a dialog with a person who is giving commands to be carried out (actually displayed on a television screen, rather than done in the physical world) and asking questions about what is going on.

To understand and carry out even a simple command, the computer must draw on a large base of programed knowledge, and a data base of information about the blocks world and the actions. It contains some subparts dealing primarily with the form of sentences--the ways in which words and phrases are combined to communicate messages. Other parts combine this information with a knowledge of word meanings to figure out what the person intends. Others use a base of information about the "blocks world" and a memory of what has been going on to figure out how to carry out requests, and to deduce the correct answers to questions.

This paper is not the appropriate place to go into the details of how the program functions or the specific theories underlying it. Rather, the emphasis is on understanding how the process of constructing such programs can be a direct method for expanding our knowledge of how human language comprehension might function in a similar environment.

Dialog with a computer program. (From Winograd, 1972.)

Person's utterances are in italics.

*How many blocks are not in the box?*

FOUR OF THEM.

*Is at least one of them narrower than the one which I told you to pick up?*

YES, THE RED CUBE.

*Can the table pick up blocks?*

NO.

*Can a pyramid be supported by a block?*

YES.

*Will you please stack up both of the red blocks and either a green cube or a pyramid?*

OK. (carries out the action)

*Which cube is sitting on the table?*

THE LARGE GREEN ONE WHICH SUPPORTS THE RED PYRAMID.

*Had you touched any pyramid before you put the green one on the little cube?*

YES, THE GREEN ONE.

*When did you pick it up?*

WHILE I WAS STACKING UP THE RED CUBE, A LARGE RED BLOCK, AND A LARGE GREEN CUBE.

*Why?*

TO GET RID OF IT.

*Why did you do that?*

TO CLEAN OFF THE RED CUBE.

*Why did you clear off that cube?*

TO PUT IT ON A LARGE GREEN CUBE.

*How many objects did you touch while you were doing it?*

FOUR OF THEM.

*What did the red cube support before you started to clear it off?*

THE GREEN PYRAMID.

*A "steeple" is a stack which contains two green cubes and a pyramid.*

I UNDERSTAND.

*Are there any steeples now?*

NO.

*Build one.*

OK.

## II. Artificial Intelligence and Cognitive Theory

Thinking and language have been studied in many ways, but the fundamentals have evaded our grasp. Anyone who looks at the history of the study of the mind will be impressed by the lack of agreement over what the proper domain of study should be, what would constitute an appropriate theory, or what type of explanation is of interest. There is no science of cognition corresponding to our understanding of physics, chemistry, or even the complex physical systems underlying life. We need to construct a theoretical underpinning--a set of metaphors and mechanisms like the billiard balls of basic Newtonian physics or the spinning shells of chemical valence.

Artificial intelligence is based on the belief that the study of computational processes is the basic viewpoint from which to build such a theory. Many aspects of our thought, our language, and our consciousness have analogs in the technical language of information and computation. AI has gone against the extreme behaviorist viewpoint that only observable stimulus and response can be talked about in a theory of mind. Just as physicists had to postulate "forces" and "atoms" and "fields" they could not directly observe, we need to form a conceptual framework of underlying objects and operations for cognition. Within this framework, we can build detailed models and test how well they mirror reality.

In the few years since people first programed computers to process symbolic information, many programs have been written to imitate human intellectual abilities in limited ways. There have been programs to play games, to do chemical analyses from spectrograms, to answer algebra problems posed in English, and many other demonstrations of what might be called "intelligence." The most important lesson has been that the computer is more than a helpful calculating tool, or a way of testing sufficiently formalized theories. It is also a metaphor for the mind, and in trying to write "intelligent" programs, it provides a basic metaphor for looking at human cognitive functions. Even if there are areas of the human mind to which this metaphor does not ultimately extend, we have good reason to believe that an information processing view can give us insight into many human mental functions, including those involved in language comprehension.

If the computational metaphor simply stimulated new ways of thinking about language comprehension, the result would be valuable, but would not constitute a separate domain of research and study. Artificial intelligence provides not just a paradigm, but also a body of theoretical and technical tools, developed in computer science and computer programing. These include programing languages like LISP for list processing, PLANNER and QA4 for manipulating propositional models, and production systems for simulating problem-solving processes. They include specific mechanisms like heuristic search of a tree of possible acts, move-generation and evaluation methods for game playing, resolution theorem proving, and the use of hierarchical context trees to represent dynamically changing situations. As each new invention is explored, researchers try to push its applicability to the limits, using it for the widest range of problems they can imagine.

Often the range of an idea cuts across the divisions imposed by an orientation to the tasks being done. A principle may apply to parts of language understanding, visual information processing, and game playing, while not being relevant to other parts of those same activities. As more research is done, we are gradually coming to understand the advantages and limitations of each of these concepts and devices. There is no magic principle which makes a mechanism intelligent. Rather, intelligence comes from having a broad flexible variety of mechanisms, and being able to apply them appropriately to the needs.

At a more theoretical level, computer science is developing formal notions of concepts like information and computational complexity. AI in particular has dealt with what is known as the "representation problem." In describing a system, whether a person, an animal, or a computer program, which is able to process and store information about the world in which it functions, we must have a theoretical framework in which to describe the mechanisms allowing it to use previously stored or built-in knowledge to process its inputs, and decide on behavior. We are concerned with developing a formalism, or "representation," with which to describe that knowledge. We seek the "atoms" and "particles" of which it is built, and the "forces" that act on it. AI has been described as the field of "experimental epistemology"--we can propose and build knowledge systems and test their consequences. Any research into language understanding will interact with research into the epistemology of commonsense knowledge and specialized knowledge domains. The examples in the introductory section were intended to point out that it is impossible to analyze what goes on in a language user without thinking about how his knowledge of the world is structured and how that structure interacts with speaking and understanding.

Current research and that of the next few years will be strongly focused on exploring different representations, in particular the effect they have on trying to build large knowledge bases. Many elegant schemes seem to break down when we try to include a domain whose size realistically approximates even a small subpart of a person's knowledge. Our explorations of representation must pay special attention to problems of modularity and interconnection that will make large systems possible, and must underlie the human reasoning system. By focusing on what is common in intelligent systems, we are working toward a general theory of representation and knowledge.

## Advantages of the AI Approach

Proponents of AI have sometimes gone overboard in the enthusiasm engendered by its new perspective. They have conveyed the impression that in a few years we have attacked and solved the problems which failed to crumble under centuries of siege by philosophers, psychologists, linguists, and logicians. This distortion feeds a strong and useful feeling of excitement and energy in its students and researchers, but to assess the degree of support to give future research, we must take a more critical comparative view.

In this section I will outline the advantages AI seems to offer in the study of language, and in the following section I will discuss some of the difficulties it poses.

Explicit goals and criteria. We have previously stated that AI research provides one clear notion of what can and should be done to study theories of cognition, and furthermore, that the test of a theory is directly operational--a program must be runnable, and in running it must simulate a part of human language behavior in some significant way. Of course, there are many problems inherent in this criterion. We would have an easy time of justifying it if we could begin with primitive elements which corresponded directly to the physical devices of the human brain, and could build a program simulating all of language behavior. But neither of these will be possible for a long time, if ever. We must start with the computational primitives we know how to build into physical systems, and we must choose some small piece of language behavior to simulate. Our simulation must attempt to duplicate some aspects of language faithfully, while treating others as secondary or irrelevant to the particular study. As in any science, the art lies in choosing the right simplifications and limitations in order to build pieces of the theory.

One of the most important aspects of an AI approach is that it gives us new ways of dividing up the problem. By trying to model language behavior in complex computer systems we can focus on the problems of coherence and system integration; on understanding what different elements go into language use and how they fit together. This emphasis is radically different from most previous work done by psychologists and linguists. In order to contrast the advantages of AI's more system-oriented view, I will discuss these earlier approaches in some detail.

*Comparison to linguistics.* In recent years there has been a great surge of interest in the problems of syntax--the rules governing the ways in which words and phrases of a language can be arranged. Chomsky and those following his views demonstrated that many of these rules could be formulated in a precise way within a framework of mathematical computation theory. Much effort has gone into specifying the exact computational formalism best suited to predict which combinations of words will be judged "grammatical" in a given language.

In a way, this approach was a great improvement over earlier approaches to linguistics, as it demonstrated the benefits of a more mentalistic approach--one which proposed formal mechanisms whose workings were not directly observable, but which had an internal consistency and coherence, and whose consequences could be shown to coincide with some aspect of actual language behavior. Linguists could begin to develop theories about what must be going on during that unobservable "flash" of language comprehension.

However, much of the early success of this approach was based on a narrow concentration on problems of syntax, avoiding issues of meaning and communication in language use. From a commonsense point of view, communication and meaning are much more central to a theory of language than the details of word ordering. But paradoxically, this centricity makes them harder to attack theoretically. They are so intertwined with our basic ability to think and reason that it becomes impossible to talk about them as a separate field of study. In order to avoid this, linguists concentrated on syntax alone, and believed that more comprehensive theories could be built later around the syntactic core.

There is much current dissatisfaction with this restrictive paradigm, and many prominent linguists are beginning to reject a syntax-dominated approach and to look at these broader issues. Recent work of Fillmore (1975), George Lakoff (1972), Robin Lakoff (1972), Halliday (1973), and many others has drawn attention to problems in describing the logical structure of meaning, and the influence of social situation on language patterns. It draws on earlier work such as Searle (1970) and Grice (1967), which started from a philosophical analysis of discourse.

The computational mechanisms developed for theories of syntax are not adequate for larger questions of meaning and the structure of discourse, and the new work just mentioned is in a descriptive style, without any coherent formalism or theoretical framework. Artificial intelligence will provide a major source of concepts and tools for this new exploration. The notions of representation and effective computation in AI have been developed to deal with issues of meaning and reasoning very much like those needed to describe what goes on in natural language discourse. Current notions like "frames" are being developed both in AI laboratories (Minsky 1975, Winograd 1975) and by linguists (Fillmore 1975) and there is rapidly accelerating convergence between the two fields.

Comparison to psychology. Just as linguists gained a precise method by limiting their attention to syntax, psychology has gone through a period of methodological precision at the expense of breadth. The focus has been on experimental paradigms in which the behavioral environment can be so constrained and simplified that a single output variable (such as response latency, accuracy of recall, or subjective measures such as similarity judgments) can be shown to vary systematically with some variation of the stimuli along simply measurable dimensions.

This effort has produced interesting results in such areas as the recall of meaning-free items (Sternberg 1969) and is perhaps applicable to some more central language areas, such as syntactic parsing (Kaplan 1975) and simple logical operations such as negation (Carpenter, Just, and Clark 1973). The results of studies in this traditional "vary the input--measure the output" mode can often be suggestive and of interest in pointing to some of the fundamental properties and operations of the system. Anyone building serious AI models must take into account the phenomena they point out. However, it seems unlikely that the kind of understanding we seek of the mind can come from a concentration on the experimental data.

The same statement holds true in a somewhat lesser degree for the more empirical areas of educational psychology, where people have made a great many observations on how various changes in the content of language material affects the way it can be understood and learned. What we need to make sense of this wealth of data, is a much more coherent picture of what is happening when a person reads or hears a sentence. Through an understanding of the interactions among knowledge, memory, and different sorts of word processing, we can begin to fit together the results derived from experiments into a framework and to apply the insights we get from that new synthesis.

Once we try to deal with the complex interactions of meaning, memory, and reasoning as illustrated in our example with Tommy, the experimental paradigm becomes impossible to apply. There is no simple, one-to-one correspondence between separate mechanisms and simple observable effects. The amount of time it takes to respond to a question, for example, is determined by an interaction of syntactic factors (the form of the sentence), reasoning (how the interpretation is affected by what the person already knows, and how an appropriate response is decided on), and a number of other issues (like the time taken by distractions resulting from associations between words in the input and other knowledge in the hearer's mind). Experiments which try to factor out these issues do so at the expense of putting people into highly artificial understanding tasks, in which their normal use of knowledge and reasoning may well be subordinated to mechanisms more suited to dealing with puzzles. There is good reason to suspect that we could not combine a parsing study, a memory study, and a logic-manipulation study to give us appropriate results for an understanding of the whole of language, if each had been done in such a way as to make the interactions as small as possible. It is a bit like studying the starter of an automobile by trying to drive down the highway without putting in any gasoline.

The interactions and organizational structures connecting different subsystems are the source of the most important properties of any complex system, whether a biological organism, a society, or a language comprehender. The AI approach, with its emphasis on building operational systems, provides strong motivation to formalize those aspects of language dependent on the interactions between all of a person's linguistic knowledge. A program cannot behave intelligently unless it deals in some way with the whole range of mechanisms underlying the understanding of what is being said. The program that produced the dialog in the introductory section of this paper placed major emphasis on the problems of representing its knowledge of meaning and language structure in a way that made it possible to combine them flexibly in understanding. Current projects in speech understanding (Woods 1974, Lesser et al. 1974, Walker 1974) are going further in trying to integrate acoustics and phonology as well.

As with linguistics, it is interesting to note that more and more psychologists are beginning to work in a style very close to AI and that we can expect the line between "psychologist" and "AI researcher" to become fainter and fainter. We can point to a good deal of work already forming the basis for a new cognitive psychology. The ideas of list processing and computation serve as a theoretical base for studies of human memory such as Sternberg (1969), Collins and Quillian (1969), Anderson and Bower (1973), and increasingly many others. Newell and Simon (1972) have used artificial intelligence models as a basis for understanding human behavior in problem-solving tasks, and others like Norman, Rumelhart, and their colleagues (1975) have been applying computational models in a variety of language behavior experiments. AI will not replace experimental psychology, because their goals are complementary, not contradictory. The interaction must be one of mutual stimulation, in which the results of experiments provoke the AI theory-builders to account for a wider range of behavior in their models, while the efforts of building intelligent programs will provide a continuing source of new potential theories and mechanisms for experimentalists to test.

## Limitations of the AI Approach

Along with the excitement and new ideas accompanying the first few years of AI research, there has also been a healthy questioning of where it is headed, and whether it is likely to succeed. In evaluating the usefulness of artificial intelligence in understanding human intelligence, there are several basic questions which must be addressed.

Ultimate limitations. Philosophers throughout the centuries have argued over the question whether man is a mechanism, accessible to the same forms of scientific explanation which serves for the brute machines we build. Some critics (e.g., Dreyfus 1972) have claimed that attempts at artificial intelligence are doomed to failure because they deal with areas demanding "uniquely human forms of information processing." There has been a good deal of rather heated debate over the detailed content of the criticisms (see Pylyshyn 1973, Papert 1968).

From the standpoint of evaluating current research, this ultimate comparison between man and machine seems the wrong question to address. The human mind is not a simple expression of a single underlying principle. It is a complex mechanism developed through eons of evolution, combining a variety of different parts, different principles, and different functions into a single integrated system. It may well be that there are aspects of emotion or perception or attention not best described in the computational metaphor of artificial intelligence. This does not in the least refute the fact that there are many ways in which the mind can be best compared to a computer. Each step in developing programs which perform in a more human intelligent way has helped to point out those aspects of thought we can analyze within a computational paradigm. Our experience in simulating human language understanding over the short history of AI leads us to believe that in this area, the amount to be gained from continued research is tremendous. We are so far from reaching any potential limitations that speculation about them is of philosophical rather than practical interest.

Testability. Artificial intelligence is at an early stage of development, with a strong emphasis on generating new ideas and possibilities for theories--even for notions of what theories can look like. We are not yet to the stage of maturity where we can devise experiments to distinguish between competing models of the large-scale structure of the mind. In fact, we are just beginning to have an idea of what those models might look like, and what ingredients they should contain. For many more years to come, the most important part of the research will be this development of new concepts, new formalisms, new kinds of theory.

Scientists trained in the empirical tradition often find it quite difficult to accept the working method which goes along with this spirit. The theories are built out of conceptual entities--programs, data, representations--which are not directly accessible to our current methods of observing the human brain. They model tasks, such as chess-playing and language understanding, for which it is extremely difficult to set up controlled experiments in which simple parameters can be varied and studied individually. Their performance is judged successful if they mimic "in essence" what a person would do in the same task, even if no details of the simulation match exactly.

These factors imply that we must set up new standards by which to judge and test models, that cannot depend on the sorts of measurements psychology has traditionally demanded. It is the responsibility of scientists working within the AI paradigm to develop standards by which to judge the validity of models, and to justify these standards to those outside the field. Indeed, this is a hard task, but it should not prevent us from exploring new techniques and reaching out for new ideas.

In setting these standards, there are some pitfalls we must avoid. In deciding whether a proposed theory "works," we must be careful not to deal with variables at too low a level of computational abstraction. Two totally different physical devices may be performing what is in essence the same computation. However, they may vary quite differently in their physical behavior--their use of space, energy, and time. Thus, if we try to apply too strict a criterion of similarity between our models and human performance using these variables, we may reject models that have captured the structure of what is going on at all but the most primitive level.

At the same time, we cannot depend on looking for naive equivalence at the level of input-output simulation. The way two different people respond to the same linguistic input, even within a simple subject area, depends on the minute details of what they know, what they have been thinking, and what linguistic patterns they personally have developed. Their answers may be essentially similar, but the test must be subjective rather than precise. In testing a model, there is no way to determine what the correct response would be at any time, although we could easily classify the great majority of possible utterances as inappropriate.

Turing (1950) proposed that we could test language models by seeing how well they could fool a person interacting with them into thinking they were human. Some work has been done (Colby 1973) on testing how well a program can mimic a specialized version of human language behavior. This "imitation" test is far from being a comprehensive answer to the problem of evaluating AI programs, but is an example of one way in which we can search for new notions of how to test our models.

Coherence. It should be clear from the discussion so far in this paper that AI is something very different from an attempt to map mind onto machine piece by piece. The major ideas are those dealing with intelligence at a more abstract level. They deal with information and process, meaning, and symbol in a way which is not specific to any one physical device, whether transistor or neuron. We are looking for those principles that describe how complex information handling systems operate, and how knowledge can be structured. The act of building a machine to perform some task does not guarantee the use of any interesting principles, nor must the brain make use of all the different methods and conceptual mechanisms we invent. But there is a coherent body of understanding we can develop from an AI viewpoint that is applicable to both mind and machine.

In a report to the British government on AI, Sir James Lighthill (1973) questioned whether there is indeed a field which appropriately combines the building of intelligent machines with understanding human intelligence. He suggested that although the building of practical intelligent devices was worthwhile for economic reasons, it had little light to shed on the human mind, and that the attempts to bridge the gap between automation and neuroscience were failures. In contradiction, it is interesting to note that there has been a good deal of synergy between different AI tasks. Ideas developed in building a program for visual scene analysis find a use in language comprehending systems, while principles of game analysis appear in programs for moving a robot about. Although only time will tell for sure, it seems highly plausible that just as there are general principles of hydrodynamics applying to both oceans and human cells, there are general principles of intelligent systems, applicable both to the human mind and to the computer programs we construct.

## The Effect of Artificial Intelligence on Education

It is amusing and enlightening to try to predict the ultimate achievements and impact of intelligent computer programs, but these speculations are often more provocative than predictive. This paper takes a more limited perspective, concentrating on the prospects for the period of funding NIE might consider--a decade or so. In the previous sections, we have looked specifically at the area of language comprehension and the ways in which current AI research can interact with more traditional disciplines to expand and restructure our theories. It is important not to see this work in isolation and, before going on to outline a proposed course of research, I want to point out some of the other potential effects of AI on education.

Practical devices. For most people, the juxtaposition of the words "education" and "computer" conjures up the image of the teaching machine. Advocates of computer-aided instruction (CAI) have embraced a wide range of schemes for putting the computer to use directly in the educational process. It has long been recognized that the ultimate effectiveness of teaching machines will be linked to the amount of intelligence embodied in the programs. That is, a more intelligent program would be better able to formulate the questions and presentations most appropriate at a given point in a teaching dialog, better equipped to understand the response of the student, and even able to analyze and model the knowledge state of the student in order to tailor the educational experience directly to that student's needs. A number of researchers (e.g., Carbonell and Collins 1973) have used the teaching dialog as a basis for looking at natural language and reasoning. This development is likely to continue and expand over the next few years.

This will be but one part of a rapidly developing sphere of practical applications of artificial intelligence. Currently, programs can perform useful services, competing with or exceeding human performance in areas such as symbolic mathematical manipulation and spectrographic analysis of chemicals. Other areas, such as medical diagnosis and automatic program writing are the focus of a good deal of new research. Information retrieval systems are benefiting from increasing natural language interaction and inference in deciding what to retrieve. We can expect our achievements at the theoretical level to go hand in hand with the development of more and more useful programs.

This paper does not deal directly with the question of computer-aided education. We hope that the research sponsored by this program will be one of the major cornerstones for improved CAI systems in the future. However, the immediate exigencies of building applied systems are quite distracting to a central theoretical orientation. The detailed practical and technical issues must (if the system is to be successful) dominate the theoretical ones. For this reason, it would be better for such systems to be developed under a program with a stronger immediate goal orientation and different criteria for choosing and judging work.

Subject-matter epistemology. Research in AI has sometimes been called "thinking about thinking," since its main concern is with the underlying principles of intelligence. But in the course of trying to write intelligent programs, we find that we are mostly engaged in "thinking about thinking about something." Whether the program is to play chess, analyze chemicals, or manipulate mathematical formulas, a tremendous amount of effort must go into structuring what we know about that domain. The programer must have taken the basic conceptual tools of AI and built an explicit representation of what a computer (or person) needs to know about the subject. This framework can be extremely useful in teaching that same subject to a person, and there is a direct educational payoff from the effort that goes into taking stock of the set of ideas in an area, organizing them, and finding their interrelations in a systematic way.

This applied epistemology will be especially important in areas such as medicine, where most of the teaching has been through experience, and where well-built structures of knowledge such as those in elementary physics and mathematics have not been developed. For a wide range of other fields as well, including much of what actually goes on in the "hard" sciences, and in areas such as effective writing, the effort of "thinking about thinking" which goes into AI research will result directly in more effective teaching of those areas.

As with CAI, there are immediate applications of these ideas which could be developed. Efforts to build models in areas like physics and medicine are going on in conjunction with educators in these fields (e.g., Schwartz 1970). Elementary school students in some experimental classes are learning skills such as arithmetic through taking the role of programer themselves, analyzing what must be put into a simple program to do the task they are learning (see Papert 1972). We can expect the number of domains to which this AI approach is applied to expand rapidly in the next few years. Although I do not propose here separate support for epistemological projects, they will form a major part of any system which purports to explain understanding.

## III. A Strategy for AI Research

The first two sections of this paper have outlined AI in general, describing its promise and the ways in which it interacts with other fields. In this section we will sketch the main directions in research being done today, and propose a policy for NIE in funding new research relevant to the goals of the essential skills program.

### Where Things Stand

Five years ago, if asked to give a description of achievements in natural language processing, an AI researcher could point to three areas of limited success. Translation projects were moribund, pushed to the limits of their approach without achieving anything better than the crudest translations. A few programs existed which achieved a much better level of language performance by limiting themselves to very restricted types of language communication, dealing with the information expressed in key words, or performing a rigidly restricted task such as solving algebra problems. Finally, there was an attempt to use "natural-language-like" systems for communicating with computers, but these made no pretense of using real language, and the user learned to speak to them, much as with any other programing language.

In the past few years, several new ideas have been developed, and we are able to write much more complex and interesting programs. Advances in representation and programing tools have made it possible to build systems (like Winograd 1972, Woods 1973) dealing with many aspects of language use, accepting facts, carrying out commands, and answering questions stated in natural language. Each system integrates many kinds of language knowledge, including grammar, semantics (how words and constructions convey meaning), and the reasoning needed to "understand" what they are being told in terms of inferences they can make in the worlds they discuss.

The performance level of these programs is impressive, and they have stirred great interest, not only among computer scientists, but in linguistics and psychology as well. They have formed the basis for speech understanding systems being designed to understand spoken requests and commands within limited domains. However, there are some fundamental issues at the heart of human language use which are still very much unexplored, and which must be at the center of research if progress is to be made. The two key problems are the structuring of large domains of knowledge and the ability to integrate flexibly multiple bodies of knowledge in the understanding process.

The AI programs of the late sixties and early seventies are much too literal. They deal with meaning as if it were a structure to be built up of the bricks and mortar provided by the words, rather than a design to be created based on the sketches and hints actually present in the input. This gives them a "brittle" character, able to deal well with tightly specified areas of meaning in an artificially formal conversation. They are correspondingly weak in dealing with natural utterances, full of bits and fragments, continual (unnoticed) metaphor, and reference to much less easily formalizable areas of knowledge.

In the past few years, we have seen the beginnings of research into these less explicit aspects of communication. Other papers in this study describe some of the work in detail, and several recent books (Schank and Colby 1973, Bobrow and Collins 1975, Norman, Rumelhart, et al. 1975) contain a large number of explorations into these new directions. Some of the research involves the development of new formalisms for grammars, involving "case" representations, in which a particular word (usually a verb) carries with it a set of expectations for other types of words or phrases that will appear in the same sentence. Others have looked into semantic representations in which a concept is represented by a structure involving a number of expected role fillers, with probable guesses used when they are not specified fully in the sentence. Others have worked at a higher level, looking at the structure of simple stories and event sequences, seeing how a program can fill in implicit chains of causation and logical connection, and use expectations to help understand the significance of the events described.

So far, all of these projects have tackled limited pieces of the communication process, and there is not yet enough theoretical background to combine them into an integrated system like the large question-answering systems of the past few years. The next few years will see the further development of these ideas into large working systems which reach a significantly higher level of performance, and are closer in many respects to the human language user.

The other main problem in existing systems is the way different areas of knowledge are integrated in an understanding process. There has been a traditional separation of linguistic knowledge into areas such as syntax, semantics, and pragmatics, and a corresponding modularity in building systems. The overall system must integrate several different pieces (such as a parser and a theorem prover) which operate quite differently and with distinct formalisms and sets of data.

Most programs have enforced this conceptual separation throughout, viewing the overall system as a set of boxes, each of which performed one part of the analysis, producing data for the next box. Some systems provide a certain amount of intermixing of these processes, allowing some pieces of the input to go through several stages before other pieces were processed, and using a small amount of feedback information. Current projects, particularly the speech understanding projects (see Newell et al. 1973, Erman 1974), are looking much more closely at this interaction, trying to build models in which knowledge of any sort can be applied whenever it seems the most certain. This is necessary for speech, since it may be necessary to understand part of the meaning of a sentence before even being able to determine fully what words are there. However, even in written language this kind of integration is vital to handling the communicative aspects of language use discussed above.

## Specific Research Support Areas

In this section I will try to define a set of possible project types for NIE support. I am intentionally not specifying specific tasks to be contracted or immediate goals to be achieved. Rather, I believe the best strategy lies in outlining a general type of support, and remaining flexible in detail. I have attempted to circumscribe a range of mutually supportive activities, and believe that it would be better to support all of them at a lesser level, rather than concentrate on only one. This, of course, interacts with the support which can be expected from other sources in one or more of these subareas. Throughout its short history, AI support has been dominated by the Advanced Research Projects Agency (Department of Defense), which has supported basic AI research of the sort described in this paper. However, it seems probable that the trend of their support will be toward more practical, more task-directed projects. The recent speech-understanding project (Newell et al. 1973) and efforts in automatic programing and automation are examples of ARPA's attempts to move towards research with more immediate practical goals, organized around principles of management aimed toward efficient directed achievement. The NIE is in a better position to support work in the more traditional academic vein of exploring a range of ideas, letting the direction be guided by the insights developed in the research, rather than a set of external milestones to be achieved. I believe that NIE support can be complementary to ARPA support along these dimensions, and that in many cases the work being done in conjunction with specific research areas might well be divided between the two. The overlap with the interests of the National Science Foundation, the National Institute of Mental Health, and other agencies is less clearcut, but should not be a major factor in defining the area to be supported since there is no other large-scale funding program currently in this general area.

I have grouped the research to be supported into four areas:

Integrated language systems
Explorations in language structure
Explorations in knowledge structures
Related linguistic and experimental studies

This classification cuts across many issues along which proposals might be divided, such as the exact goals of the system (translation, story understanding, cooperative dialog, etc.) and the domain of knowledge. I intentionally do not divide things along those lines, since the relevance of a project to the development of the field is not tied to those decisions.

The specific goals a project sets out to achieve are important to shaping its focus, but often can be a comparatively superficial way of defining the research. A system that sets out to translate and one that simply absorbs story information into a computer knowledge base are quite different in outward aim. This often corresponds to a difference in theoretical approach, but not in a clear one-to-one way. The same theoretical base might be used for both tasks, or two different ones for seemingly similar tasks. A project which sets out to build a "practical working system" will make different trade-offs in how it allocates its resources, but may be based on essentially the same ideas as another project in the form of a gedanken experiment.

Similarly, it is often misleading to judge the relevance of a project by its subject domain. A system about airline reservations may have much more in common with one being used in medical history-taking than it has with one working on office schedules. As with the choice of immediate goals, we cannot completely separate the domain subject from the sort of competence necessary and the particular organizations for knowledge which will be suitable. Researchers spend long agonizing hours trying to choose domains whose characteristics will support rather than hinder an attempt to deal directly with the most important underlying issues. In deciding on research support for projects bearing on natural language, the concern with the domain area should be almost exclusively at this level--looking for domains that encourage the asking of significant questions, rather than applying external criteria of relevance to specific educational goals.

It should be apparent from the discussion in the first two sections that the most important dimension in evaluating any AI research project is the one of theoretical gains in understanding. We are engaged in understanding how knowledge can be formally represented and manipulated, and the relevance of any piece of work to our ultimate goals must be judged on its contribution to our set of theoretical concepts and ability to apply them. This dimension cuts sharply across all of the others. In our discussion of the "payoffs" from AI we distinguished between domain epistemology and language modeling. We pointed to different sorts of applications, and a corresponding difference in research emphasis. However, in the next few years these two goals are both best served by a common concentration on underlying issues of representation and the integration of large knowledge systems.

The research sponsored under this program should remain focused on issues related to language comprehension, but the boundary should not be drawn too narrowly. AI research in the epistemology of knowledge domains will be of central relevance even when it does not address itself directly to problems of natural language communication.

Integrated language systems. This paper has emphasized the way in which problems of size and complexity play a crucial role in our understanding of language comprehension. There is a fundamental difficulty in defining a small independent research project which can attack these problems. In order to understand what is happening in systems involving large bases of knowledge and many different processes, we must work with such systems. But no one person can build all the necessary pieces, and our current level of sophistication in program writing makes it extremely difficult to integrate pieces without prior concern for their compatibility. It therefore seems a useful strategy to support some small number of projects that attempt to combine the efforts of many people into building a single comprehensive system. Within this framework, individual researchers can choose to concentrate on specific aspects, with a commitment to work in a way compatible with the system as a whole and concerned with the interactions with other parts. There are few examples of such systems currently, although the Lindsay-Norman-Rumelhart (LNR) group (Norman and Rumelhart 1975) and Schank's group (Schank 1975) have many of the desired characteristics. Such a group would involve two or three senior investigators and a few nonstudent research staffers, and could form the focus for a number of theses and dissertations.

In the future it will be possible for such a group to operate through the use of external computer resources, perhaps through a network like the ARPANET supported by the Advanced Research Projects Agency of the Department of Defense. For the next few years it would seem more profitable for such a group to be in control of its own dedicated computer resources (like the current AI groups). This means a fairly high level of financial commitment and an expectation of continuing support for a basic period on the order of 5 years, in order for the initial tool-building stages to pay off in good research. Every effort should be made to make use of existing computer systems rather than putting resources into developing new ones.

Explorations in language structure. One of the most important contributions of many projects lies in the set of tools they develop. Even though we must eventually deal with issues of complexity and integration, there is still much room for research on a smaller scale that tackles individual problems and develops specific tools and concepts. It is still the case today that most of the major advances have been made in the scope of a single doctoral dissertation, often unrelated to any other ongoing research. Many of the issues in the structure of language can be profitably attacked at this level. A project whose general approach and scope seem inadequate to the task of full language comprehension may still be of great value in developing a particular tool which other people can integrate into more comprehensive systems. Programs can be written to perform some particular piece of analysis (some examples might be syntactic parsing, the application of discourse rules to understanding a dialog, or the representation of different types of description and reference). Care should be taken to look for these advances and find ways of making use of them, rather than trying to push them (or dispute them) as models for the entire language facility. Some support should go to small projects to explore the usefulness of a particular set of methods in some corner of the vast language-comprehension process, and that other work should try to integrate the insights gained by a variety of such explorations. Projects of this sort could be supported on a 2-year basis on the scale of a single senior investigator with a small number of graduate students.

Explorations in knowledge structures. These projects would be similar in size and scope to those labeled "explorations in language structure" but would be oriented toward more general problems of representing knowledge. We have emphasized the importance of this endeavor for our ultimate understanding of language comprehension, and I believe that explicitly supporting it under the same aegis as more immediately language-oriented work may well increase the interaction and benefits to both areas. The choice of domain should be evaluated with an eye toward its language potentialities. Work in areas such as the modeling of intuitive commonsense physics, medical diagnosis, and the knowledge of programs and programing itself might be included under this subarea.

Related studies in linguistics and psychology. As pointed out in Section II, there is much to be gained from an interaction between the ideas being developed in AI and the traditional concerns and methods of linguists and experimental psychologists. Although I do not see this style of crossdisciplinary study as the major thrust of the proposed research effort, it seems that a small amount of support could provide quite beneficial results. Projects supported at the same level as in the previous two paragraphs (actually somewhat less, since computer resources need not be involved) could be carried out by people interested in making use of AI models, or finding the connections between them and other work. The fields of linguistics and psychology have developed their own funding patterns, and this program should not attempt to take over any large quantity of it. However, there is something to be gained by the increased visibility and interchange of ideas which comes from having the most related of the research in these areas under the same roof.

## Areas Not Mentioned for Support

In describing potential areas for support, I have left out a number of possible others. Some may have just been forgotten, and should be included. Others I have considered and left out for specific reasons:

Speech comprehension and production. There is much to be learned in this area and it is closely related to many of our basic goals. I have excluded it partly because its emphasis tends to be somewhat towards the more technical acoustic issues, and partly because there is currently a large-scale effort being made by ARPA in the area, and it might be wise to wait for the results before deciding on the direction for future research.

Interactions between language and visual processing (reading). As with speech, this area is not excluded for lack of interest, but because its concerns and methods are different enough from the more central issues of knowledge and language structure that trying to include it would make the program too diffuse.

Physiological and neurological studies. It will be many years before the understanding of the human nervous system will reach a point where it can be mapped in a significant way onto cognitive processes as abstract as language understanding. Although this remains a tantalizing long-range goal, I believe there is little to be gained from prematurely trying to combine efforts of such different methodological prejudices.

# References

Abelson, R. P. The structure of belief systems. In R. C. Schank and K. M. Colby (eds.), Computer models of thought and language. San Francisco: Freeman, 1973.

Anderson, J. and Bower, G. Human Associative Memory. Washington, D.C.: Winston-Wiley, 1973.

Bobrow, D. G. and Collins, A. M. (eds.). Representation and Understanding. New York: Academic Press, 1975.

Bobrow, D. G. and Raphael, B. New programming languages for artificial intelligence research. Computing Surveys. 1974, 6(3), 153-174.

Carbonell, J. R. AI in CAI: An artificial intelligence approach to computer-aided instruction. IEEE Transactions on Man-Machine Systems. 1970, MMS-11, 190-202.

Carbonell, J. R. and Collins, A. M. Natural semantics in artificial intelligence. Reprinted from American Journal of Computational Linguistics. 1974, I, Mfc. 3.

Carpenter, T., Just, M., and Clark, H. On the meeting of semantics and perception. In W. Chase (ed.), Visual Information Processing. New York: Academic Press, 1973.

Charniak, E. Towards a model of children's story comprehension. (AI TR-266). Cambridge: Massachusetts Insitute of Technology, 1972.

Colby, K. M. Simulations of Belief Systems. In R. C. Schank and K. M. Colby (eds.), Computer models of thought and language. San Francisco: Freeman, 1973.

Collins, A. M. and Loftus, E. F. A spreading activation theory of semantic processing. Psychological Review. 1975, 82(4).

Collins, A. M. and Quillian, M. R. Retrieval time from semantic memory. Journal of Verbal Learning and Verbal Behavior. 1969, 8.

Dreyfus, H. L. What computers can't do: a critique of artificial reason. New York: Harper and Row, 1972.

Erman, E. (ed.). IEEE Symposium on Speech Recognition. Computer Science Department, Carnegie-Mellon Univ., April 1974.

Fillmore, C. The case for case. In Bach and Harms (eds.), Universals in Linguistic Theory. 1968, 1-90.

Fillmore, C. Against a checklist theory of meaning. Proceedings of the First Annual Meeting of the Berkeley Linguistics Society. Institute of Human Learning, Berkeley, 1975.

Goldman, N. and Riesbeck, C. A conceptually based Sentence Paraphraser (Stanford AI Memo 196). Stanford, California: Stanford Univ., 1973.

Gordon, D. and Lakoff, G. Conversational postulates. Papers from Seventh Regional Meeting. Chicago Linguistic Society. Chicago: Univ. of Chicago Linguistics Department, 1971.

Grice, H. P. Logic and conversation. Unpublished lecture notes from William James Lectures at Harvard, 1967.

Halliday, M. A. K. Explorations in Language. 1973.

Hewitt, C., Bishop, P., and Steiger, R. A universal modular ACTOR formalism for artificial intelligence. Proceedings of the Third International Joint Conference on Artificial Intelligence. 1973, 235-245.

Kaplan, R. On process models for sentence analysis. In D. A. Norman, et al., Explorations in cognition. San Francisco: Freeman, 1975.

Lakoff, G. Hedges: A study in meaning criteria and the logic of fuzzy concepts. Papers from the Eighth Regional Meeting, Chicago Linguistics Society. Chicago: Univ. of Chicago Linguistics Department, 1972.

Lakoff, R. Language in context. Language. December 1972, 907-927.

Lesser, V. et al. Organization of the HEARSAY II speech understanding system. IEEE Symposium on Speech Recognition. Computer Science Department, Carnegie Mellon Univ., April 1974, 11-21.

Lighthill, Sir James. Artificial intelligence: A general survey. Artificial Intelligence. Science Research Council, London, April 1973.

Minsky, M. (ed.). Semantic Information Processing. Cambridge: MIT Press, 1966.

Minsky, M. A framework for representing knowledge. In Winston, P. (ed.), The psychology of computer vision. New York: McGraw-Hill, 1975.

Minsky, M. and Papert, S. Progress Report. Massachusetts Institute of Technology, Artificial Intelligence Laboratory, 1972.

Moore, J. and Newell, A. How can MERLIN understand? In Gregg (ed.), Knowledge and Cognition. Lawrence Erlbaum Associates, 1973.

Newell, A. and Simon, H. A. Human Problem Solving. Englewood Cliffs: Prentice-Hall, 1972.

Newell, A. et al. Speech Understanding Systems. Amsterdam: North-Holland, 1973.

Norman, D. A. and Bobrow, D. G. On data-limited and resource-limited processes (CSL Report 74-2). Palo Alto, California: Xerox Palo Alto Research Center, May 1974.

Norman, D. A., Rumelhart, D. E., and the LNR Research Group. Explorations in cognition. San Francisco: Freeman, 1975.

Papert, S. The Artificial Intelligence of Hubert L. Dreyfus: A Budget of Fallacies (MIT AI-Memo-154). Cambridge: MIT, January 1968.

Papert, S. Teaching children thinking. Mathematics Teaching. The Bulletin of the Association of Mathematics, 58, Spring 1972.

Pylyshyn, Z. W. Minds, machines and phenomenology: Some reflections on Dreyfus' "What computers can't do." Cognition. 1973.

Quillian, M. R. Semantic memory. In M. Minsky (ed.), Semantic Information Processing. Cambridge: MIT Press, 1968.

Rieger, C. Conceptual memory (AIM-233). Stanford, California: Stanford Univ., 1974. In R. Schank (ed.). Conceptual information processing. Amsterdam: North-Holland, 1975.

Riesbeck, C. K. Computational understanding: Analysis of sentences and context. Unpublished doctoral dissertation. Stanford Univ., 1974. In R. Schank (ed.), Conceptual information processing. Amsterdam: North-Holland, 1975.

Schank, R. C. (ed.). Conceptual information processing. Amsterdam: North-Holland, 1975.

Schank, R. and Colby, K. (eds.). Computer Models of Thought and Language. San Francisco: Freeman, 1973.

Schank, R. C. Conceptual dependency: A theory of natural language understanding. Cognitive Psychology. 1972, 3, 552-631.

Schank, R. Identification of conceptualizations underlying natural language. In R. Schank and K. M. Colby (eds.), Computer Models of Thought and Language. San Francisco: Freeman, 1973.

Schwartz, W. B. Medicine and the computer: The promise and problems of change. The New England Journal of Medicine. 283, 23, December 1970.

Searle, J. R. Speech Acts. Cambridge, 1970.

Simmons, R. F. Semantic networks: Their computation and use for understanding English sentences. In R. C. Schank and K. M. Colby (eds.). Computer models of thought and language. San Francisco: Freeman, 1973.

Sternberg, S. Memory-scanning: Mental processes revealed by reaction-time experiments. American Scientist, 57, 4, 1969.

Turing, A. M. Computing machinery and the mind. Mind: A Quarterly Review of Psychology and Philosophy, 1950.

Walker, D. E. The SRI speech understanding system. IEEE Symposium on Speech Recognition. Computer Science Department, Carnegie-Mellon Univ., April 1974, 32-37.

Winograd, T. Understanding Natural Language. New York: Academic Press, 1972.

Winograd, T. When will Computers Understand People? Psychology Today. 7:12, May 1974.

Winograd, T. Frames and the declarative-procedural controversy. In D. G. Bobrow and A. M. Collins (eds.). Representation and Understanding. New York: Academic Press, 1975.

Winograd, T. Five lectures on artificial intelligence (Stanford AI-Memo-246). Stanford, California: Stanford Univ., September 1974.

Winston, P. H. (ed.). The psychology of computer vision. New York: McGraw-Hill, 1975.

Woods, W. A. Transition network grammars for natural language analysis. Communications of the ACM. 1970, 13(10), 591-606.

Woods, W. A. Motivation and overview of BBN SPEECHLIS: An experimental prototype for speech understanding research. Proceeding of the IEEE Symposium on Speech Recognition. Carnegie-Mellon Univ., April 1974.

Woods, W. A. Progress in natural language understanding: An application to lunar geology. AFIPS Conference Proceedings. 42, 1973, National Computer Conference and Exposition.

# ARTIFICIAL INTELLIGENCE, LANGUAGE, AND THE STUDY OF KNOWLEDGE

Ira Goldstein
Seymour Papert
Marvin Minsky

Artificial Intelligence Laboratory
Massachusetts Institute of Technology

July 1975
Revised October 1975

This paper studies the relationship of artificial intelligence to the study of language and the representation of the underlying knowledge supporting the comprehension process. It develops the view that intelligence is based on the ability to use large amounts of diverse kinds of knowledge in procedural ways, rather than on the possession of a few general and uniform principles. We conclude with a discussion of the radical impact artificial intelligence may have on education if the principles it uses to explore the representation and use of knowledge are made available for students to use in their own learning experiences.

The research and writing of this paper was sponsored under NIE Contract 400-75-0012. The views and conclusions contained are those of the authors and should not be interpreted as representing the policies, either expressed or implied, of the National Institute of Education or the United States Government.

# TABLE OF CONTENTS

CHAPTER 1--INTRODUCTION 29
  Artificial Intelligence and the Study of Language
  Artificial Intelligence as the Study of Intelligent Processes
  A Knowledge-Oriented Rather than Power-Oriented Theory of Intelligence
  Artificial Intelligence as a Procedural Theory of Knowledge
  Organization of This Paper
  Educational Applications

CHAPTER 2--FRAMES AS A UNIFYING APPROACH TO REPRESENTING KNOWLEDGE 34
  Case Grammars
  Understanding a Sentence
  The Procedural Idea of a Frame-Keeper
  A Formal Realization of the Frame Idea
  Frames for Representing Knowledge About Nouns
  Extending the Frame Idea to a Theory of Context
  Frames for Thematic Knowledge

CHAPTER 3--THE STRUCTURE OF AN AI THEORY 45
  Issues in Computational Epistemology
  Representing Declarative Knowledge
  Representing Procedural Knowledge
  Representing Knowledge in Highly Structured Packets
  The Demon Solution to the Reference Problem
  The Virtue of Particular Knowledge

CHAPTER 4--SELF-KNOWLEDGE 50
  Kinds of Self-Knowledge
  The SHRDLU Program and the Blocks World
  Answering Elementary Why Questions
  Self-Reference in Grammatical Systems
  Debugging Knowledge

CHAPTER 5--CONTROL STRUCTURES 56
  Controlling Grammatical Computations
  Augmented Transition Networks
  An Issue for the Future--Deterministic Parsing

CHAPTER 6--THE CLASSICAL, ROMANTIC, AND MODERN PERIODS OF AI 63
  A Change in Perspective
  Some Unanswered Technical Issues in Frame Theory

CHAPTER 7--THE IMPACT OF ARTIFICIAL INTELLIGENCE ON EDUCATION 66
  If the Machines Get Smarter, Will the Children Get Smarter?
  Intelligent Computer-Aided Instruction
  The White Box Approach to Language Studies
  The Articulate Learner
  Epistemology and Education

EPILOG 73
  Controversies About Educability and Innateness

BIBLIOGRAPHY 74

# INTRODUCTION

## Artificial Intelligence and the Study of Language

This essay is about the impact on the study of language of a new style of thinking about cognition known as Artificial Intelligence (AI). The goals of the essay are: to convey a sense of the intellectual personality of artificial intelligence, and to identify some trends in its development which we see as especially relevant to the study of language and of the role of an AI approach to language in education.

## Artificial Intelligence as the Study of Intelligent Processes

Before we embark on the substance of this essay, it is worthwhile to clarify a potential source of confusion. For many, AI is identified as a narrowly focused field directed toward the goal of programing computers in such a fashion that they acquire the appearance of intelligence. Thus, it may seem paradoxical that researchers in the field have anything to say about the structure of human language or related issues in education. However, the description is misleading. It correctly delineates the major method of the science, that is, the use of computers to build precise models of cognitive theories. But it mistakenly identifies this activity as the only purpose of the field. Although much practical good can come of more intelligent machines, the fundamental theoretical goal of the discipline is understanding intelligent processes independent of their particular physical realization.

Naturally, understanding intelligence abstractly has intimate interactions with understanding it in its natural form. Thus, AI and psychology have close ties. Indeed, AI has been called "theoretical psychology" (Newell 1973). AI draws from psychology a set of basic concerns--understanding language, perception, memory, problem solving. Psychology, in turn, acquires a new framework of computational ideas for expressing cognitive theories. Indeed, a major purpose of psychology becomes the discovery not of the class of programs that could possibly explain a given intelligent behavior (this is the archetypical concern of AI), but rather the particular program that a given individual actually possesses.

## A Knowledge-Oriented Rather than Power-Oriented Theory of Intelligence

Given the general perception of AI as the computational study of intelligence, it still remains to identify the particular set of questions and theoretical constructs that a researcher in the field brings to bear on cognitive problems. In the earliest years of the field (i.e., during the late fifties and early sixties), the hope was that researchers could identify a few simple but very powerful techniques that together could be used to create intelligent programs. During this period, theorem proving and search played a primary role. Indeed, to many AI has become identified as the science of heuristic search, that is, the study of procedural techniques for exploring state spaces too large to be explored in an exhaustive fashion. During these early years, an AI paper on language might have begun by representing the state space by means of a formal grammar and then have focused on the issue of achieving successful parsings in reasonable times (Green 1963; Lindsay 1963).

Today there has been a shift in paradigm. The fundamental problem of understanding intelligence is not the identification of a few powerful techniques, but rather the question of how to represent large amounts of knowledge in a fashion that permits their effective use and interaction. This shift is based on a decade of experience with programs that relied on uniform search or logistic techniques that proved to be hopelessly inefficient when faced with complex problems embedded in large knowledge spaces. The solution from the current point of view is for the problem solver (whether man or machine) to know explicitly how to use its knowledge--with general techniques supplemented by domain specific pragmatic knowhow. Thus, we see AI as having shifted from a power-based strategy for achieving intelligence to a knowledge-based approach. Minsky and Papert characterize these two points of view as follows:

> The POWER strategy seeks a generalized increase in computational power. It may look toward new kinds of computers ("parallel" or "fuzzy" or "associative" or whatever) or it may look toward extensions of deductive generality, or information retrieval, or search algorithms--things like better "resolution" methods, better methods for exploring trees and nets, hash-coded triplets, etc. In each case the improvement sought is intended to be "uniform"--independent of the particular data base.
>
> The KNOWLEDGE strategy sees progress as coming from better ways to express, recognize, and use diverse and particular forms of knowledge. This theory sees the problem as epistemological rather than as a matter of computational power or mathematical generality. It supposes, for example, that when a scientist solves a new problem, he engages a highly organized structure of especially appropriate facts, models, analogies, planning mechanisms, self-discipline procedures, etc. To be sure, he also engages "general" problem-solving schemata but it is by no means obvious that very smart people are that way directly because of the superior power of their general methods--as compared with average people. Indirectly, perhaps, but that is another matter: a very intelligent person might be that way because of specific local features of his knowledge-organizing knowledge rather than because of global qualities of his "thinking" which, except for the effects of his self-applied knowledge, might be little different from a child's.
>
> (Minsky-Papert, 1974. P. 59)

It is worthwhile to observe here that the goals of a knowledge-based approach to AI are closely akin to those which motivated Piaget to call his research center "Centre d'Epistemologie Genetique." Or more precisely, to his reason for calling himself an "epistemologist" rather than a psychologist. The common theme is the view that the process of intelligence is determined by the knowledge held by the subject. The deep and primary questions are to understand the operations and data structures involved. Neither Piaget nor the authors consider the physical (biological or electrical) mechanisms underlying them as the source of intelligence in any structural sense (although intelligence could not happen without some such embodiment).

Perhaps the knowledge paradigm is most clearly characterized by a developmental perspective. Naturally, intelligence is impossible without a certain basic "hardware" capability. But our position and Piaget's is that this capability is already present in the very young child. The subsequent development of intelligence is not due to the emergence of new biological capabilities in the individual. Rather, increased performance is due to the acquisition of knowledge. This knowledge can range from specific facts about a domain to general problem-solving schemata. Indeed, it may even be knowledge about how to learn (for example, facts about debugging techniques). Some facts result only in local improvements in the individual's capabilities, while others may have global effects, giving rise to a stage transition in Piagetian terms. From the computer standpoint, we see the need for some improvements in speed and memory capabilities, especially for such special purpose applications as vision and speech. But the fundamental difficulties facing researchers in the field today are not limitations of hardware, but rather questions about how to represent large amounts of knowledge in ways that still allow the effective use of individual facts.

## Artificial Intelligence as a Procedural Theory of Knowledge

The reader may perhaps wonder at this point how AI is to have an identity of its own if the knowledge strategy is pursued, i.e., if it asks only for explicit representations of the knowledge of particular domains. Isn't this simply the concern of the specialized field of study associated with that domain? How does this notion of AI, for example, distinguish it from linguistics when the domain is that of language competence?

The answer lies in identifying two concerns of AI that are uniquely procedural in nature. The first is that the knowledge of a particular domain must be represented in a form that is usable. A linguist may content himself with the specification of some attribute of grammatical knowledge by defining the appropriate transformation: An AI researcher must go one step further by asking how this transformation can be woven into the total web of linguistic information in such a way that the comprehension or generation process is possible in reasonable computational times. Thus, AI brings to the traditional study of a given field the quest for identifying pragmatic knowledge, namely knowledge of "how" traditional facts are to be used.

There is a second concern distinguishing AI from traditional studies of a given domain: its concern for identifying the manner in which competence in different domains interacts. Again drawing upon an example from the study of language, a grammarian can legitimately focus on attributes of grammar divorced from the semantics or use of language. The essential concern of AI, however, is to construct frameworks in which all these kinds of knowledge can successfully interact in order to allow the complete comprehension process to take place. Thus, both concerns unique to AI arise from its responsibility to structure knowledge into procedural systems that can actually solve problems.

## Organization of This Paper

Our goal in the remainder of this paper will be to provide the reader with insight into the structure of theories currently being formulated in AI within this knowledge-based paradigm. Chapters 2 and 3 illustrate the AI approach by discussing a particular example--frame theory--which many people in the field are currently exploring and find quite exciting. Chapter 2 introduces the subject by approaching frame theory as an AI extension of the traditional linguistic notion of case grammars from a narrowly syntactic theory to a general representation for knowledge. Chapter 3 discusses the various attributes of AI theories that frames exemplify. In chapters 4 and 5, we deepen our view of AI as a computational style of epistemology by discussing two aspects of cognition and language to which AI has made unique contributions--the role of self-knowledge and the importance of control structure. Chapter 6 concludes our theoretical discussion by briefly viewing in perspective the intellectual history of the field.

These topics do not by any means exhaust the contribution of AI to the study of language. But we believe they provide an excellent illustration of the AI method. And that seems to us the only serious purpose to be served by a paper of this length that is accessible both to students of the field and to nonspecialists.

Throughout this paper, we avoid extensive excursions into formalism and jargon. More seriously, we also ignore, or pass over in a cursory way, a variety of serious technical issues, in order to focus attention on the overall enterprise of a computational investigation of knowledge. We feel especially justified because of the tendency in the literature to focus on narrow technical issues. Nevertheless, to compensate for this broad focus, we provide in chapter 6 a list of the technical questions we feel are critical for the reader who wishes to pursue this point of view more deeply. Also, the reader should recognize that the research we point to as illustrating our position varies in many important aspects. However, we feel that the critical need at this time is to develop a feeling for the paradigm which characterizes AI as a method for investigating knowledge and language, and it is to this goal that we direct our efforts. Finally, we and others in the AI community expect to develop further the ideas raised in this paper. The interested reader is invited to write us for references to the latest research in this area.

## Educational Applications

The final part of the essay (chapter 7) will touch schematically on some of the consequences for education of artificial intelligence research on natural language. These consequences are of very different sorts. Some of them, no doubt for most people the most obvious, are consequences of actually possessing machines capable of understanding what a student says. Obviously such machines would greatly improve the quality of computer-aided instruction (in the sense of tutorial programs, programed instruction administered by computers, drill and practice, computerized advice, etc.).

Much deeper consequences, however, are related to the intellectual content of artificial intelligence research on natural language. As a new kind of cognitive theory it has obvious implications for education. It is our personal opinion that these consequences will very deeply transform the theory and practice of education. Although we will not try to argue for this position at length, we shall try to convey the flavor of the perceptions on which it is based.

# FRAMES AS A UNIFYING APPROACH TO REPRESENTING KNOWLEDGE

## Case Grammars

Our goal in this chapter will be to illustrate AI's generalization of the powerful linguistic notion of case grammars to support a theory of knowledge allowing the representation of procedural knowledge for constructing a parse, as well as the expression of related semantic, contextual, and thematic facts. In fact the frame theory which we shall discuss draws upon more than purely linguistic considerations. In the original formulation by Minsky (1975), the notions of frames and frame systems were developed in equal measure from problems arising from the construction of adequate theories for vision, for memory, for logical reasoning, and for the comprehension of language. However, we shall narrow our focus in this chapter to consider the utility of frames solely in relation to linguistic knowledge. In the next chapter, we step back from the realm of language to consider a variety of larger issues that characterize the AI approach to representing knowledge.

The case frame approach to linguistics as developed by Fillmore (1968) and others constrains the inquiry to the realm of syntactic knowledge. Within this sphere, the basic idea is to organize structural knowledge of the sentence around the verb. The result is a <u>case grammar</u> which specifies for each verb the cases it requires. As an illustration, the following phrase structure grammar specifies the underlying case structures developed by Celce-Murcia (1972). Simmons (1973) used them in his language comprehension and generation programs.

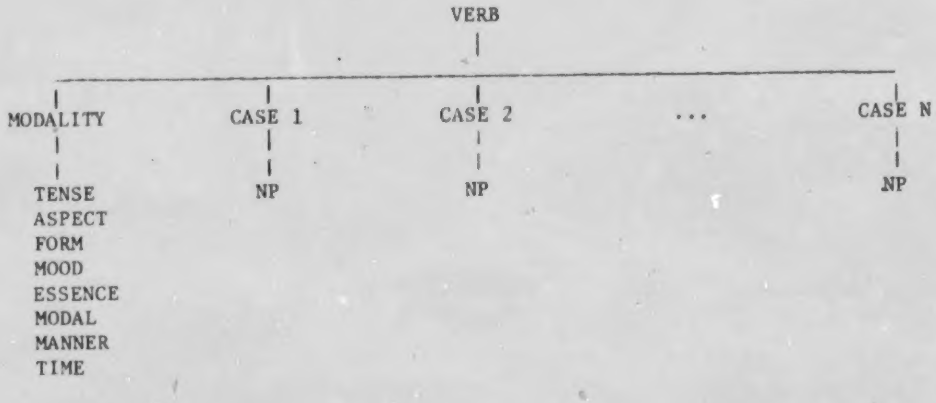
```
S => MODALITY + PROPOSITION
MODALITY => TENSE, ASPECT, FORM, MOOD, MODAL MANNER, TIME
PROPOSITION => VERB + (CASEARGUMENT)*
VERB => run, walk, break, etc.
CASEARGUMENT => CASERELATION + [NP/S]
NP =>(PREP) + (DET) + (ADJ)* + (N) + N + (S/NP)
CASERELATION => causalactant, theme, locus, source, goal
```

(Simmons 1973. P. 68)

Figure 1 gives the general syntactic form for the case structure of a sentence. To illustrate, let us use the traditional linguistic device of choosing a representative sentence around which to organize our observations. Hence, S1:

(S1) John tickled the girl with the feather.

For S1, a traditional phrase-structured grammar might produce the parse illustrated by Figure 2, while a case-oriented approach would represent the sentence as in Figure 3. Figure 3A shows the case relations for this verb following Celce-Murcia, while Figure 3B follows the formulation of Chafe (1970).


VERB
MODALITY
CASE 1
CASE 2
...
CASE N
NP
NP
NP
TENSE
ASPECT
FORM
MOOD
ESSENCE
MODAL
MANNER
TIME

The cases with their intuitive meanings:

| | |
|---|---|
| CAUSALACTANT1: | Actor, manipulator, causal agent |
| THEME: | Object |
| CAUSALACTANT2: | Instrument, second causal agent |
| LOCUS: | Location or path of activity |
| SOURCE: | Origin |
| GOAL: | Destination |

The modalities:

| | |
|---|---|
| TENSE: | Present, past, future |
| ASPECT: | Perfect, imperfect |
| FORM: | Simple, emphatic, progressive |
| MOOD: | Declarative, interrogative, imperative |
| ESSENCE: | Positive, negative, indeterminate |
| MODAL: | May, can, must |
| MANNER: | Adverbial |
| TIME: | Adverbial |

Figure 1

Case Structure Following Celce-Murcia from Simmons (1973)

(S1) JOHN TICKLED THE GIRL WITH A FEATHER.

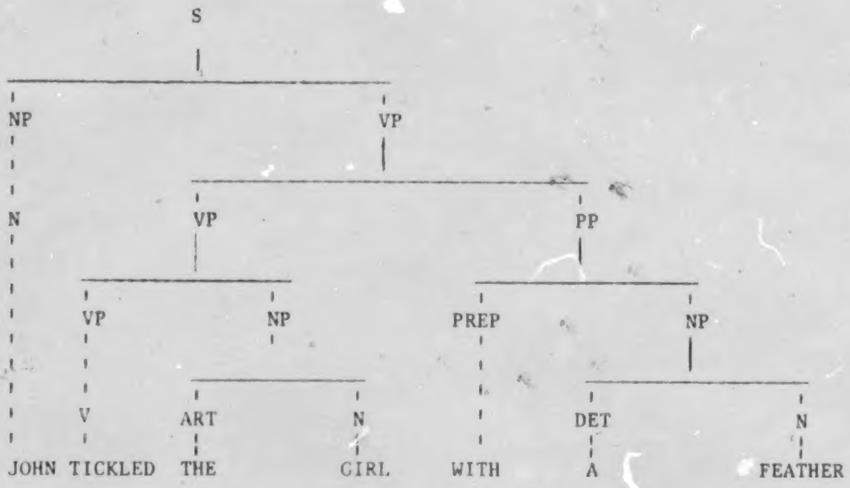

S
NP
VP
N
VP
PP
VP
NP
PREP
NP
V
ART
N
DET
N
JOHN
TICKLED
THE
GIRL
WITH
A
FEATHER

Figure 2

A Phrase-Structured Parse of S1

| | | |
|---|---|---|
| TICKLE: | CAUSALACTANT1 | John |
| | THEME | the girl |
| | CAUSALACTANT2 | a feather. |

Figure 3A

A Case-Structured Parse of S1 Following Celce-Murcia

| | | |
|---|---|---|
| TICKLE: | AGENT | John |
| | PATIENT | the girl |
| | INSTRUMENT | a feather. |

Figure 3B

A Case-Structured Parse of S1 Following Chafe

## Understanding a Sentence

To make the linguistic enterprise tractable, the linguist deliberately ignores the issue of understanding an utterance beyond the narrow concern of identifying its syntactic structure. But AI is forced to address this larger issue in order to design computational models that actually act appropriately in response to the utterance. Thus, AI is willing to consider the senses of understanding S1 illustrated by the following three listeners:

> A 5-year-old child, who has a clear image of John using the feather to evoke giggles.
>
> A slightly more sophisticated listener capable of taking account of a context such as: "All the girls in the room had strange hats. One was decorated with a large feather, another with a towering confection of artificial fruit. John tickled the girl with the feather."
>
> A linguist who understands the sentence as an example of a well-known ambiguity in parsing.

It is obvious that all the subjects use some grammatical knowledge. For example, the order of the words tells them that it was John who did the tickling and not the girl. But grammatical knowledge alone is not sufficient even to account for the syntactic parsing of the sentence, i.e., the interpretation of the phrase "with the feather" as adjectival or adverbial. This point is made clear by contrasting the sentence with S2:

(S2) John tickled the girl with the funny nose.

Even our 5-year-old would not be likely to read this sentence as if John used the nose as a tickling instrument. The difference is not puzzling to commonsense: After all the child knows that feathers are good for tickling, while noses are more likely to be used for smelling or as personal identifiers. But the difference should trouble the cognitive theorist, who knows how to describe syntactic structures in fine detail but who has no better description than naive commonsense of the child's knowledge about feathers and noses. And the difference is relentlessly present for the computer-oriented theorist who absolutely needs to find ways to formalize it.

## The Procedural Idea of a Frame-Keeper

Let us begin our generalization of case grammars by discussing informally how a procedural case theory might be organized. In the next section, we will discuss work by Riesbeck (1974) that represents one formal realization of this point of view.

In our informal minitheory there is associated with each important word in the sentence an active agent, strictly defined as a computational procedure. We might just as well anthropomorphize the agent (for the moment anyway) as a person. Let's call them Frame-Keepers, for reasons which will soon become apparent. The process of understanding the sentence will consist of an interaction between these frame-keepers wherein each tries to fulfill certain needs. For example, the needs of the tickle frame-keeper include:

> Find the agent (the person responsible for the tickling)!
> Find the patient (the person being tickled)!
> Find the instrument (the means by which the agent is tickling the patient)!

Thus, the tickle frame-keeper is actively seeking to fill these slots written into the definition. If all of them can do so to their satisfaction (an idea we will pursue shortly) the sentence is to be considered as parsed.

The essence of the AI extension of the case frame idea is that the frame for the verb contains not only syntactic information about the expected location of the underlying cases, e.g., in a declarative sentence that the subject is the agent, the object the patient, and the "with" prepositional phrase the instrument, but also semantic and pragmatic facts. Thus, we imagine that for our tickle frame, the frame-keeper has access to world knowledge that states that feathers are typical tickling instruments, while other kinds of objects, for example, dumbbells, are not. Hence, it would be the tickle frame-keeper who would be in a position to reject (for S2) that the "with" phrase refers to the tickle instrument, knowing that one generally does not tickle with noses.

In the next section, we discuss an existing program that associates both syntactic and semantic knowledge with individual words in order to comprehend the sentence, and we indicate how to extend the frame idea to nouns. Given that extension, the process of comprehending S2 becomes clearer: Not only does the tickle frame reject the nose phrase, but the girl frame accepts it as a legitimate descriptor of people. Thus, the system has two sources of evidence, both supporting one another regarding the intended parse.

## A Formal Realization of the Frame Idea

Our discussion of the tickle frame and frame-keepers represents the kind of concern for active knowledge that AI brings to linguistic studies. C. Riesbeck (1974, 1975) has designed and implemented a program that, although not frame-based, shares many important properties with the approach outlined above. Riesbeck's parser maps natural language into the conceptual dependency structures designed by Schank (1973). S3 is typical of the sentences that Riesbeck's system analyzes.

(S3) John gave Mary a beating.

Procedural information is associated with each word in the sentence, triggered when the parser encounters the word in its left-to-right analysis. This information constructs an instantiated case-framelike structure. It also contains expectations serving to clarify multiple word senses and properly decode the syntactic cues that indicate the fillers for each slot in the case-frame associated with the verb. Figure 4 gives a trace of the analysis of the parser.

The system, called MARGIE, of which Riesbeck's parser was a part, exhibited its understanding by generating paraphrases of the original sentence and by

| STEP | WORD READ | REQUESTS WAITING | REQUESTS TRIGGERED | ACTIONS TAKEN |
|---|---|---|---|---|
| 0 | none | 1-is there an NP? | none | none |
| 1 | John | 1-is there an NP? | 1 | assume "John" is subject of the verb to follow |
| 2 | gave | 2-is the current NP a human?<br>3-is the current NP an object?<br>4-is the current NP an action?<br>5-true | 5 | assume the word "to," if it appears, introduces the recipient of the "giving" |
| 3 | Mary | 2-is the current NP a human?<br>3-is the current NP an object?<br>4-is the current NP an action? | 2 | assume Mary is the recipient of the "giving" |
| 4 | a | 3-is the current NP an object?<br>4-is the current NP an action?<br>6-true | 6 | save the current list of requests and replace it with:<br>7-does the current word end an NP? |
| 5 | beating | 7-does the current word end an NP? | none | none |
| 6 | period | 7-does the current word end an NP? | 7 | build the NP "a beating" and reset the list of requests |
| 7 | none | 3-is the current NP an object?<br>4-is the current NP an action? | 4 | assume the NP action is the main action of the clause, the subject (John) is the actor and the recipient (Mary) is the object |

Figure 4

Trace of the Riesbeck Parser on "John gave Mary a beating."

From (Riesbeck 1974. P. 31)

making deductions based on it (Schank 1972). The basic operation was to comprehend the sentence by mapping it into the conceptual dependency representation developed by Schank (1973). Figure 5 illustrates the conceptual dependency diagram generated for the above sentence. The details of Schank's representation are not important here, except to note the manner in which the system has understood that the sentence refers to a series of actions by John, involving the motion of his hand. Riesbeck's parser has not erroneously interpreted "give" to mean that John is transferring the possession of something to Mary.

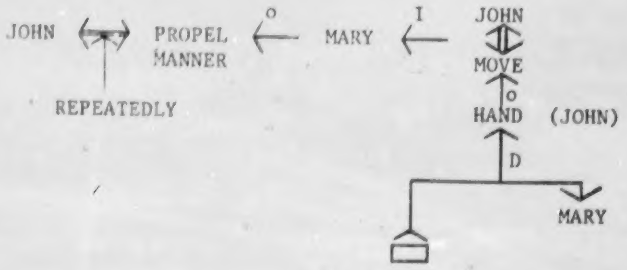


Figure 5

Conceptual Dependency Representation of the meaning of the sentence: "John gave Mary a beating."

From (Riesbeck 1974. P. 108).

## Frames for Representing Knowledge About Nouns

The case-frame idea in its original linguistic form was oriented towards representing knowledge about verbs. AI, pursuing the goal of comprehensive theories of knowledge, has generalized the idea to allow frames for nouns.

Consider again the sentence, "John tickled the girl with the feather." We imagine that part of the knowledge for understanding this sentence properly is that the verb-frame for tickle has as a typical instrument "feathers." Such verb-centered knowledge is supplemented by frames for such nouns as "girl" which contain pointers to subframes describing typical appearances, purposes, activities, etc. The appearance subframes would indicate that girls do not "have" feathers. They might under closer scrutiny indicate that "feathers" are a possible but unlikely kind of dress. However, the strong default for "feathers" as a typical tickling instrument would predominate.

In the previous paragraph, we wrote as if the tickle frame could have direct, explicit knowledge about feathers. A fundamental and pervasive class of problems concerns the relative feasibility of achieving this as compared to various schemes for indirect reference to feathers via some super-set or description. For example, the tickle frame might know that "long, soft objects" are good instruments for tickling and somehow use this as a means of finding a connection to feathers. Most early workers in AI (like most contemporary psychologists) were strongly averse to the direct approach and so a great deal of attention has been given to the invention of "search" or "information retrieval" methods to mediate the indirect approach. Current thinking (especially among frame-theorists) leans toward very much more direct knowledge than was previously considered feasible. This places a heavy demand on memory, and less on retrieval mechanisms. This appears to some to be "heavy-handed." But we do know that human memory has a very large capacity when appropriately used and no one has been able to propose very sophisticated retrieval mechanisms. So perhaps it is nature (rather than the theorists) who has adopted the "heavy-handed" solution.

A related example shows where noun-frame information might predominate. Consider the three sentences:

(1) I frightened the dog with the loud noise.

(2) I frightened the dog with the brown fur.

(3) I frightened the dog with the loud bark.

We believe most readers would understand (1) by assuming that the "loud noise" was the instrument of "frighten," because of both the verb-frame knowledge that loud noises often can be used to frighten and the noun-frame knowledge associated with dogs, indicating that "noise" is not "possessed" by them. On the other hand, in sentence (2), the noun-frame knowledge for dogs would surely indicate that they typically have "fur"; hence the phrase "with the brown fur" would be understood to be modifying the noun "dog." In sentence (3), we suspect that our population of readers would be divided, depending on whether their frame for "frightening" contained the typical instrument of emitting a loud noise, describable as a bark, with a stronger emphasis than their frame for "dogs" contained the default property of emitting noise by barking. (This extension of frames to nouns has not yet reached the point of development in the form of programs. However, a recent thesis by C. Drake (1975) extends the details of the theory beyond the outline provided by Minsky to a form that is closer to the implementation state.)

## Extending the Frame Idea to a Theory of Context

Procedural frames not restricted to verbs are powerful tools for dealing with the set of problems traditionally called "context." Indeed, once we break away from the verb-centered sentence as the unit for analysis, it becomes possible to talk about "scope" in a possibly very large text. This too is a topic of current research and there are few programs implemented as yet. One of the systems that does exist, however, is Schank's successor to the MARGIE system that generates paraphrases of simple stories (Schank 1975). The essential theoretical idea in the new system is the existence of frames (called scripts by Schank) that describe typical activities in terms of his conceptual dependency primitives. An example is the following informal sketch of a frame for going to a restaurant:

Script: restaurant
Role: customer, waitress, chef, cashier
Reason: to get food so as to go down in hunger and up in pleasure.

Scene 1 Entering
PTRANS--go into restaurant
MBUILD--find table
PTRANS--go to table
MOVE--sit down

Scene 2 Ordering
ATRANS--receive menu
ATTEND--look at it
MBUILD--decide on order
MTRANS--tell order to waitress

Scene 3 Eating
ATRANS--receive food
INGEST--eat food

Scene 4 Exiting
MTRANS--ask for check
ATRANS--give tip to waitress
PTRANS--go to cashier
ATRANS--give money to cashier
PTRANS--go out of restaurant

(Schank 1975. P. 117)

The image of understanding a story is mapping the sentences of the story into the actions described in the frame. Unstated facts described in the frame are assumed. Thus, mapping a story in which the sentence appeared: "John left a tip" into this frame would allow the system to know that the "tip" was left for the waitress, a fact contained in the frame but not in the story.

## Frames for Thematic Knowledge

Our previous examples have discussed extensions of the case-frame idea to incorporate into the frame a larger fraction of the knowledge associated with comprehending a sentence than is included within the purely grammatical realm. We would like to conclude this chapter on frames as a theory of knowledge by extending our representation beyond the sentential level to include the kinds of knowledge that link sentences together. This idea represents the frontier of current research; we will have few formal examples and have to content ourselves with some observations on the direction in which we expect this research to go.

To develop this notion of thematic frames, let us consider the problem of understanding a fairytale. Our goal is to represent the knowledge that links sentences in a story together into a coherent whole. To begin, we will first examine some work by Rumelhart (1975), which addresses this question, but employs a phrase-structured grammar (augmented by semantic interpretation rules) to represent the structure of the story. Typical of the rules that his system possesses are:

RULE 1:STORY → SETTING + EPISODE.

The setting is a statement of the time and place of the story as well as an introduction for the main characters, e.g., "Once upon a time, in a far away land, there lived a good king."

RULE 2:SETTING → (STATIVES)*

Episodes consist of events followed by reactions.

RULE 3:EPISODE → EVENT + REACTION

An event is the most general category of the grammar.

RULE 4:EVENT → EPISODE|CHANGE-OF-STATE|ACTION
|EVENT+EVENT

A reaction consists of an internal and an overt response.

RULE 5:REACTION → INTERNAL-RESPONSE + OVERT-RESPONSE.

There are a large number of internal-responses. Two are:

RULE 6:INTERNAL-RESPONSE → EMOTION|DESIRE.

Overt-responses can be either an action or a series of attempts.

RULE 7:OVERT-RESPONSE → ACTION|(ATTEMPT)*

Attempts also have an internal structure.

RULE 8:ATTEMPT → PLAN + APPLICATION.

The plan is developed to accomplish a desire. The application is the actual execution of the plan.

For each of these rules, there are semantic interpretation rules that indicate, for example, that Rule 1 means that both the setting and the episode must be present, while Rule 3 (with the same syntactic "+" connective) indicates that the event is the cause of the reaction and precedes the reaction in time.

Let us now ask how the skeleton represented by this story grammar might be extended into a frame system. The basic change is the addition of large amounts of highly structured and particular knowledge. This information, which we shall call *default knowledge*, includes information about the various aspects of the story, e.g., setting, episodes, and internal and external responses. Indeed, the proposal is that the system knows a few "favorite" stories in extensive detail. Thus, with only a brief introduction to the story such as "Once upon a time and long ago," listeners are able to bring to mind an elaborate setting corresponding to their primary defaults--perhaps the Camelot setting, the Black Forest setting, or the humble cottage, depending on the idiosyncracies of the frame system of the individual listener. Of course, conflicts will arise from differences between the story and the default, causing one of two results: either exceptions are entered into the evolving data base representing the system's comprehension of the story, or, if the conflict is too radical, the entire default setting is replaced with another. In either case, the result is that far more detail than is actually stated in the story is assumed by the listener through this mechanism of default knowledge.

Just as there can be default knowledge about states such as setting, so too can there be default knowledge about the expected sequence of episodes in the story, where each episode is itself a frame (e.g., a frame for the initial state--a damsel in distress, followed by a sequence of frames describing a typical rescue operation by a handsome prince and hindered by a wicked witch). This organization of frames into state action sequences is essentially Schank's script idea, discussed earlier.

Default knowledge about expected themes, morals, narrative structures may also be included. We expect, for example, for a fable to have a moral, for the prince and princess to live happily ever after in a children's fairytale (though not in the original primitive folklore).

The point of this discussion is to articulate this model of comprehension in which the process is not one of literally understanding the text, but instead is more one in which the text triggers rich, highly structured knowledge packets that supplement the literal content, provide expectations for the remainder of the story, and place the story in a context of related knowledge. Thus, understanding is essentially a process of evoking and then debugging existing knowledge packets.

# THE STRUCTURE OF AN AI THEORY

## Issues in Computational Epistemology

At this point, we have seen how the linguistic theory of case grammars has been extended by AI to become a knowledge-oriented theory of language. In this chapter, we step back from a purely linguistic orientation in order to gain a perspective on the general theory of intelligence and knowledge embedded in this point of view. Specifically, we shall examine:

(1) The manner in which frames exhibit a readiness to use new, flexible, and complex data types;

(2) The fashion in which frames represent a synthesis of procedural and declarative knowledge;

(3) The sense in which the use of an anthropomorphic metaphor has become a technical device in AI theory;

(4) The trend toward the representation of knowledge in larger, more structured units (which frames exemplify);

(5) A concern for allowing interactions between diverse kinds of knowledge;

(6) A desire for expertly controlled processes.

## Representing Declarative Knowledge

A central concern of AI is to construct new data types and to explore their capabilities. The frame idea is still evolving in AI, and so it is difficult to give a precise definition. Minsky gives the following description of a frame in his original paper on the subject:

> We can think of a frame as a network of nodes and relations. The "top levels" of a frame are fixed, and represent things that are always true about the supposed situation. The lower levels have many terminals--"slots" that must be filled by specific instances or data. Each terminal can specify conditions its assignments must meet. (The assignments themselves are usually smaller "sub-frames.") Simple conditions are specified by markers that might require a terminal assignment to be a person, an object of sufficient value, or a pointer to a sub-frame of a certain type. More complex conditions can specify relations among the things assigned to several terminals.
>
> (Minsky 1975. P. 212)

We see from the above description that the traditional declarative data structures of objects, properties, and relations play an important role in the definition of a frame. This aspect of frame theory has its roots in earlier work in AI on the representation of knowledge as attribute-value pairs (e.g., the BASEBALL program (Green 1963)) and as semantic nets (Quillian 1968). The concept of a semantic net has already had an important impact on cognitive psychology in the work of Anderson and Bower (1973), and Norman (1975). Frame theory generalizes the concept of a semantic net by adding the following ideas:

(1) The organization of these relations, not into a uniform set, but into a highly structured set of contexts (the frames themselves);

(2) The inclusion of defaults for properties and relations representing the expected values and supplying common-sense knowledge about the context (possibly not supplied by the actual text); and

(3) The infusion of procedural knowledge, which we shall discuss below.

## Representing Procedural Knowledge

In the preceding chapter, we saw informally how procedural expertise is part of a frame. This expertise took the form of a frame-keeper, a procedure whose action is to examine the current context for data that can fill the slots of the frame. To clarify the concept further, we begin with a gedanken example of how an AI program might emulate a certain aspect of the way a mathematician thinks. The example goes right into the domain of the logician and proposes a very different way to cope with the existence of paradoxes, such as the Burali-Forte paradox in mathematics and such paradoxes in ordinary language as the use of self-reflexive assertions such as "All Cretans are liars" (spoken by a Cretan and meaning: Cretans never tell the truth). The traditional logician's recourse is to seek an axiom from which no such contradictions can be deduced. Our theory of what a working mathematician actually does is to use principles of reasoning that could easily lead to paradoxes if they were to be applied in that direction. But the mathematician, aware of this danger, refrains from using them in directions known to be dangerous. It is as if he posted a sentry near dangerous places in his intellectual terrain. The job of the sentry is to issue warnings if any reasoning process comes too close to the danger spot. In AI we see our job to be developing theories in which the concept of such a sentry can be easily formalized.

The anthropomorphic idea of posting a sentry is typical, we maintain, of the way in which sensible people have thought intuitively about this kind of problem. The contribution of computer science is to show how to translate this intuitive matter into a technical and well-defined form. If one represents a knowledge system as a set of propositions, the idea of introducing the sentry is outrageous, or at best a literary conceit. But if we represent the knowledge system in the first place as a set of interactive processes, the sentry is merely one more process. Naturally one has to be sure that this new process is well defined, but that is a technical problem. The conceptual change is that it is now perfectly natural and coherent to introduce such entities.

From this perspective, a fundamental contribution of AI to epistemology is made clear; and that is, the systematic introduction of active agents into epistemological theory constructions, so that, for example, an item of knowledge, a concept, or whatever, is seen as an active agent rather than a passive manipulatable object.

The gedanken example from mathematics seems to us well suited to illustrate the force of this concept in an intuitive way. But there are real applications applied to real problems. A well-known example is Charniak's method for dealing with reference problems in children's stories. To illustrate, consider the following story fragment from Charniak (1972).

> Today was Jack's birthday. Penny and Janet went to the store. They were going to get presents. Janet decided to get a kite. "Don't do that," said Penny. "Jack has a kite. He will make you take it back."

The goal is to construct a theory that explains how the reader understands that "it" refers to the new kite, not the one Jack already owns. Purely syntactic criteria (such as assigning the referent of "it" to the last mentioned noun) are clearly inadequate, as the result would be mistakenly to understand the last sentence of the story as meaning that Jack will make Janet take back the kite he already owns.

Resolving reference is a traditional problem of linguistics that cannot be handled by purely syntactic techniques. Our solution will involve the realization of our intuitive frame-keeper as a precise procedural entity, the demon. This is a productionlike entity that is activated by some aspect of the state of the computational world within which it is defined. But we will postpone the details of this solution until the next section in order to first give some perspective on the underlying AI philosophy.

## Representing Knowledge in Highly Structured Packets

The fundamental frame assumption is the thesis that fewer situations than we think are really as new as we think. Most situations in which people find themselves have sufficient in common with previously encountered situations for the salient features to be preanalyzed and stored in a situation-specific form. Whenever we encounter these situations, we are able to use predigested, highly particular knowledge. This explanation is in sharp contrast with the views of intelligence propounded by theorists who would prefer to believe that our minds work (and have to work!) in a deductive fashion from very general principles.

For our kite example, it is clear that one cannot know that "it" refers to the new kite without knowledge about the trading habits of our society. One could imagine a different world in which newly bought objects are never returned to the store, but old ones are. The question we raise here is how this knowledge might be represented, stored, and made available to the process of understanding Charniak's story. To see the spectrum of possible theories, consider two extremes.

An extreme in the frame direction. The listener has stored away in long-term memory a frame for birthday parties in which certain events and problems are explicitly represented. These include the problem of choosing a present with quite explicit reference to the three or four "possible dangers to be thought about" including "recipient won't like it and will be disappointed" and "recipient already has one and will return it."

An extreme in the antiframe direction. From a strictly logical point of view it is not necessary to have a preconstructed party frame in order to assign the proper referent to "it." It is unnecessary to have been to a party or to have given a present. It would be sufficient to have stored away for general use items of information such as:

(1) {(x has been bought at store y) and (x is in good condition)}

implies (store y will exchange x for z)

(2) {(x is new) and (x has not been used)}

implies (x is in good condition).

The advantages of this method (if it is workable) are apparent: less memory space, easier updating, modularity, etc. The major disadvantage is that we then have to face the question of how these two items of information are chosen from the presumably very large stock of knowledge, presumably including many other items rather similar to these. How does the interpretive system know which items to use? An even more serious problem takes the form: How does the system know that any appeal to knowledge about the world is necessary? Are we to assume that simple syntactic rules (e.g., based on word order and proximity) are never used? But if they are sometimes used, what restrains them in this case?

By providing possible answers to these questions, the frame theory shows itself to be much more than a method for information retrieval or heuristic search. One such answer is that Charniak's reference problem simply does not arise, in the sense that the Birthday-Present-Frame-Keeper will already have assigned the new kite to a slot for "returned gifts" before the pronoun is processed!

## The Demon Solution to the Reference Problem

Charniak's 1972 research on constructing a model for comprehending children's stories predates the articulation of the frame idea, but presages important notions of that concept in several ways. In particular, his formal realization of a frame was in the form of base knowledge about a variety of situations that arise in the context of these stories. The mechanism of his program was for the content of sentences to evoke this base knowledge with the following effect: Demons were created to monitor the possible occurrence in later sentences of likely (but not inevitable) consequences of the given situation. Thus, for our story fragment the birthday knowledge creates expectations about the need for participants at the party to buy presents and the possible consequence of having to return these gifts. Hence, these demons expect the possibility of Jack already possessing the present and the resulting need for Janet to return it, where it is known to be the present.

It is interesting to note that we cannot apply this demon about presents unless one makes several commonsense inferences. For example, the demon might be a formal representation of the following fact:

> If a person P1 does not want a present X, then he may request the giver P2 to return the present.

To apply the demon, the antecedent--i.e., that P1 does not want the present X--must be triggered by an assertion in the data base being created to represent the meaning of the story. (One may think of the data base as containing the kernel sentences of the story with each word mapped into its proper sense and reference problems clarified.) But the story does not include explicitly all important facts. Look back at the story. Some readers will be surprised to note that the text itself does not state (a) that the presents bought by Penny and Janet were for Jack, (b) that the kite bought by Janet was intended as a present, and (c) that having an object implies that one does not want another. All of the above facts are inserted into the data base by other demons activated by the birthday frame. Thus, Charniak's notion of a demon is particularly satisfying because demons not only handle reference problems, they are also used to insert various reasons, motivations, and purposes not explicitly stated in the story data base.

In the current view of frame theory, this insertion of unsaid information assumes a paramount role. From this point of view, it is not accidental that so much is left out of the story. Rather the typical situation in comprehension is to be faced with a set of clues that evoke a rich and detailed knowledge structure--the frame--that supplies the unstated details. Naturally, these defaults may be inappropriate for some situations and, in those cases, the text must supply the exceptions. But in general, reference, deduction, and comprehension are made possible only by knowing a great deal about the situation being discussed.

## The Virtue of Particular Knowledge

AI is a young science and the ideas that we have discussed in previous sections are under constant development, elaboration, and change. However, it has been our goal to give the reader a feeling for the kind of enterprise that AI is rather than a final statement on its theoretical form. A central point has been the way that many kinds of knowledge play a role in cognitive processes. In particular we considered for the reference problem the need for real world as opposed to syntactic knowledge and, more profoundly, particular as opposed to general knowledge. A major point of our analysis, i.e., of the knowledge-as opposed-to-power theory of intelligence, is that particular knowledge can often do the job for which general knowledge is traditionally thought to be necessary.

It is encouraging that there are parallel issues in the theoretical writings of Piaget and the Geneva School of genetic epistemologists. In particular Piaget's description of the knowledge of the young ("preoperational") child is much closer to the frame-specific knowledge paradigm than to the generalist power paradigm.

## SELF-KNOWLEDGE

In the next two chapters, we leave the context of frame theory in order to broaden our insight into the notion of AI as a new kind of epistemology. We shall discuss two concerns central to AI theories that rarely appear in more traditional formulations of knowledge and intelligence. The first is the formal representation and use of self-knowledge, and the second is the importance of control structures.

### Kinds of Self-Knowledge

As we shall see, an interest in self-knowledge will not take us outside our concern for understanding the comprehension process. Indeed, we shall find that this kind of knowledge plays an obvious role in allowing the system (whether computer model or individual) to answer questions about its own performance as well as a role in the less obvious situation of handling special cases that arise in parsing the syntactic structure of the utterance.

There are many kinds of self-knowledge. Perhaps the simplest is the ability to observe one's own behavior, allowing the individual or program to answer elementary questions about its actions. A more skillful use of self-knowledge would involve being able to take account of these observations and direct one's problem solving accordingly. Still a more sophisticated program might, in certain circumstances, employ this self-observation to learn by debugging faults in its procedures, generalizing its knowledge, or perhaps suitably specializing it to form more efficient techniques for particular problems. We shall illustrate each of these kinds of self-referential abilities below in relation to the Blocks World, a favorite miniworld for AI programs in which a one-armed robot is asked to build various structures from blocks placed on a table (Figure 1).

### The SHRDLU Program and the Blocks World

Winograd's SHRDLU program represented a tour de force of design and an important benchmark in computer comprehension of natural language. In this section, we introduce the program briefly and in the next two sections illustrate several ways in which it manifests self-knowledge.

SHRDLU is a program that converted such sentences as "Pick up the big red block and put it in the box" into commands for a blocks world problem solver. The success of its comprehension was demonstrated by the program successively carrying out the instructions in a simulated world displayed on a cathode ray screen which contained a one-armed robot, a table, a box, and a collection of cubes and pyramids.

The program is capable of comprehending sentences that reached the degree of complexity illustrated by the following examples:

THE ROBOT'S WORLD

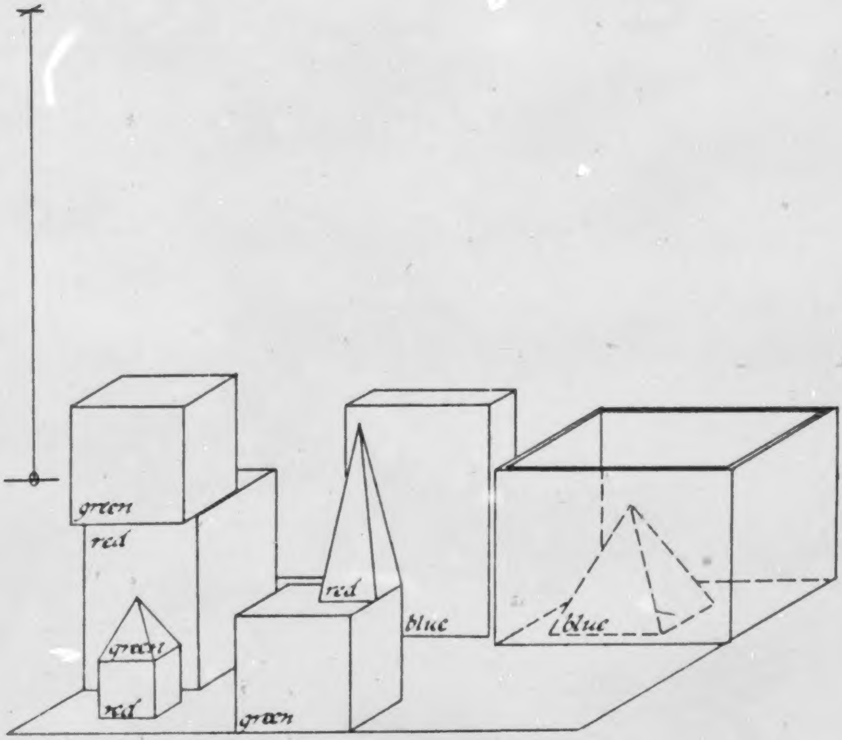

green
red
green
red
red
green
blue
blue

Figure 6

"Pick up the big red block."

From (Winograd 1972. P. 8).

(1) Find the block which is taller than the one you are holding and put it into the box.

(2) Will you please stack up both of the red blocks and either a green cube or a pyramid.

(3) Does the shortest thing the tallest pyramid's support supports support anything green?

In SHRDLU, there are a multitude of different knowledge structures. For example, the system maintains event-history lists recording each event, why it occurred, and what it accomplished; numerical coordinate data structure recording the dimension and location of objects; a symbolic data base in which assertions about the state of the world, about global knowledge, and about active procedures are stored; and a Fodor-and-Katz-style marker hierarchy (Katz 1964) indicating the relation between words. But, perhaps most significantly, a major part of the systems knowledge was represented directly as procedures, thereby achieving far more flexibility and power than more traditional "declarative" representations. These different knowledge packets are used to answer questions, to perform actions, to update its knowledge, and so forth.

## Answering Elementary Why Questions

An elementary example of self-knowledge is SHRDLU's capacity to generate names for assertions in its data base, events, noun groups, and other entities. With these names, the system can refer to these entities, associating them with times, places, and other information, thus allowing the system to answer elementary kinds of questions about its behavior. Figure 7 illustrates this ability.

## Self-Reference in Grammatical Systems

SHRDLU also illustrates a second level of self-referential ability in which the system can not only answer questions, but make decisions based on self-knowledge. This is the manner in which the system has access to the current state of the parse, and to the past history of the parse, illustrated by the fashion in which Winograd's program parses conjunctions.

To parse sentences with words like "and," the system interrupts upon detecting the conjunction, notes the last constituent formed, and attempts to find a duplicate constituent following the conjunction. If unsuccessful, it looks further back in the history of the parse until it finds a substructure that symmetrically appears following the conjunction. A similar capability for self-reference is found in Woods Augmented Transition Network approach to parsing (Woods 1970). Elements of this approach are discussed below in the section on controlling grammatical computations.

It is interesting to note that "*and*" is not handled simply by augmenting the basic systemic grammar of SHRDLU. By some standards, this moving outside the formalism of one theoretical mechanism would be unaesthetic by being contrary to Occam's Razor. But experience has shown that the management of diverse knowledge packages is the rule rather than the exception in the

23. had you touched any pyramid before you put the green one on the little cube?

YES, THE GREEN ONE.

The system deduces from its memory what event is being referred to, and can relate other events to it in time. The analysis includes complex tenses of verb groups, and there are special facilities for keeping track of the various forms of irregular verbs like "have."

24. when did you pick it up?

WHILE I WAS STACKING UP THE RED CUBE, A LARGE RED BLOCK AND A LARGE GREEN CUBE.

A time is described by naming the main goal which was being worked on.

25. why?

The system remembers immediately previous sentences to understand questions like this and the following ones.

TO GET RID OF IT.

By keeping track of selected parts of the original subgoal tree, the system has some understanding of its own motives.

26. why did you do that?

TO CLEAN OFF THE RED CUBE.

"That" refers to the action mentioned in the answer.

27. why did you clear off that cube?

Here, "that" means the one just referred to.

TO PUT IT ON A LARGE GREEN CUBE.

28. why did you do that?

TO STACK UP THE RED CUBE, A LARGE RED BLOCK AND A LARGE GREEN CUBE.

29. why did you do that?

BECAUSE YOU ASKED ME TO.

It remembers what it did, not how the request was worded.

Figure 7

Examples of SHRDLU's ability to answer WHY questions.

From (Winograd 1972. P. 13).

architecture of complex problem-solving systems. The obligation is simply one of clearly defining the interactions between these knowledge packets.

## Debugging Knowledge

Recently the Blocks World has served as a domain for investigating self-knowledge of a deeper kind. This has been in the form of a system called HACKER (Sussman 1974), capable of learning to solve block construction problems of the complexity that SHRDLU can handle. (Remember that for SHRDLU, the problem-solving procedures are preprogramed and not learned.) The learning is achieved through HACKER possessing an expertise in debugging. It is our belief that debugging knowledge is absolutely fundamental to the acquisition and operation of any program as complex as human thought and language.

Although this work was not done specifically as part of a natural language comprehension system, it provides the basic understanding to enable such a system to answer questions about why a particular method was used, in a far deeper fashion than SHRDLU's elementary abilities allowed. Knowledge of one's own problem-solving strategies supports an understanding that could discuss both the chronology of events and the underlying plans and their evolution through succeeding stages of generalization and debugging.

To illustrate, we shall develop an example from Sussman. HACKER initially might be compared to a small child interested in constructing block structures, but not very knowledgeable about the problems involved. HACKER begins with the ability to place one block on top of another (by executing the primitive PUTON) and with an elementary planning capability. A typical elementary goal is to construct a tower from three blocks, expressed formally to the program by the conjunction (GOAL(AND(PUTON A B)(PUTON B C))). HACKER attempts to solve this problem in a "linear" fashion first by planning to achieve (PUTON A B), and then as an independent second goal (PUTON B C), producing the sequence of block states illustrated below:

```
              Trace of (ACHIEVE (AND (ON A B) (ON B C))

state 1:  A    B    C

               A
state 2:  B    C      ;PUTON A B has just been executed.

state 3:  A    B    C    ;The top of B has been cleared as a required
               ;prerequisite of PUTON B C.  HACKER is one-handed
               ;and must lift blocks from the top.

               B
state 4:  A    C      ;PUTON B C has just been completed.
```

Obviously a bug has occurred. The action of clearing the top of B as a prerequisite to achieving (ON B C) resulting in state 3 has undone the first subgoal (ON A B), the purpose of which was to place block A on block B.

One solution to this difficulty which AI programs of the past might have attempted (and which children occasionally pursue), is to return to the original state and try some different construction strategy, without attempting to learn anything from the current failure to guide the next attempt. HACKER contains a new kind of knowledge, namely knowledge of how to debug. This debugging knowledge knows that linear solutions are only first approximations and that sometimes interactions between the subgoals of a conjunction must be taken into account. In this particular case, HACKER recognizes that the prerequisite of one subgoal CLEARTOP conflicts with the purpose of a brother goal. In such situations, it applies the debugging strategy of seeking a reordering of the brother goals such that the goal with the prerequisite is achieved first. In this particular problem, HACKER

(1) Recognizes that the bug is due to the general bug type of "prerequisite clobbers brother goal."

(2) Applies the debugging strategy of reordering the goals.

(3) Edits in long-term memory the procedure for constructing towers of three blocks to take account of this ordering requirement.

(4) Records a comment to itself for subsequent problems of this kind that states that when building block towers of any size, ON chains should be ordered from the bottom up.

Thus, HACKER exhibits expertise about debugging and about repair, coupled with the ability to examine its own problem-solving goals and actions so that it is able to apply this debugging expertise to its own reasoning. It represents a model of skill acquisition whose competence is based on self-knowledge about its behavior and meta-knowledge about the nature of plans and bugs.

We emphasize the importance of debugging skill because a vital element of intelligence is knowing how to handle a variety of situations. We cannot know all about these situations, because each real world situation is different. What we must know is how to tell which old situation is sufficiently like the present one, and how to adapt--"debug"--the old procedure for handling the old problem to a new procedure to deal with the new situation. Thus, subjects like analogy, similarity, metaphor are at the heart of the new formalisms. In the older methods of AI--linguistics, logic, and psychology--these subjects were embarrassing, hard-to-explain phenomena. In the new, debugging technology, they become concrete, manipulable types of knowledge.

Finally, without a theory of debugging, our image of frame-based comprehension would have a fatal flaw. Understanding by using large packets of previously known information requires the ability to debug these knowledge structures for the particular application at hand, if the system is to be able to deal with novel situations.

## CONTROL STRUCTURES

We have seen the way in which AI epistemology is opening new areas of study in its concern for self-knowledge and, in particular, debugging skill. A second dimension along which AI's approach to epistemology is opening new horizons and fundamentally differs from older approaches is its attention to control structure. Indeed the idea of control structure could scarcely exist until one begins to see knowledge as an aggregate of active agents needing control.

### Controlling Grammatical Computations

"First-order" theories, such as those composed of generative grammars, are constructed under strong assumptions of modularity and independence. These assumptions are heuristic ones that enable the epistemological analysis to get underway, but include an expectation of a subsequent need to study the interactions. Controlling computation by choosing between alternative chunks of knowledge for relevance, deciding when a given approach should be terminated, and generally guiding the reasoning process are among the important kinds of interactions which we must eventually understand.

When context-free grammars were first applied to the problem of comprehending text (rather than only generating grammatical sentences), people immediately recognized that more structure was necessary if the parsing times were to be kept within reasonable bounds. The essential problem is to provide information for choosing whether one should apply a given rule. This basic concern characterizes the basic approach of both Augmented Phrase Structured Grammars (Heidorn 1975) and Augmented Transition Nets (Woods 1970). The ATN approach has served as the basis of several important language programs and we shall discuss it further below. See, for example, Simmons (1973) in which ATN's are used both for parsing into an internal representation and for generation of English output from this representation.

### Augmented Transition Networks

The Augmented Transition Network formalism for expressing grammars, developed by Woods, was designed to extend the power of finite state recognition grammars. His first step was to generalize the finite state graph to allow recursion; thus essentially obtaining the power of a context-free grammar while still preserving the perspicuity of a network representation. A recursive call is indicated by the label of an arc being a state name rather than a terminal. This directs the recognizer to jump to that state, run until acceptance, and then return control to the state pointed to by the arc. Essentially a pushdown store has been provided to keep track of the bookkeeping involved in allowing recursion. Figure 1 illustrates a recursive transition net. An example of recursion occurs in the interpreter's transition from state S to state Q1 by means of the recognition of a noun phrase.

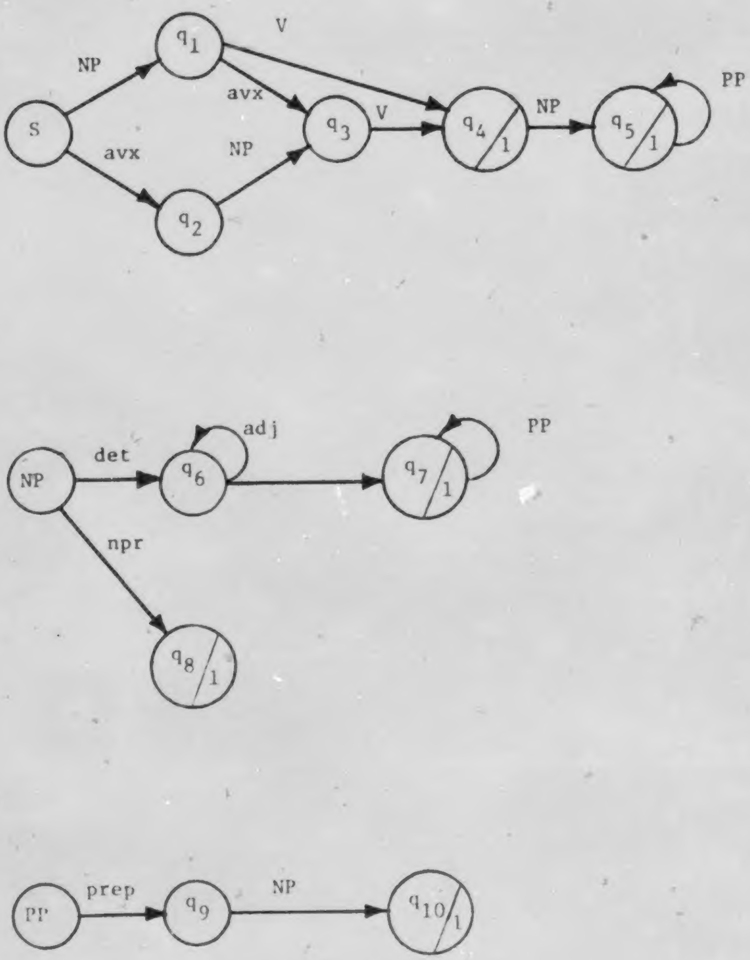

V
NP
q1
avx
S
q3
V
q4/1
NP
q5/1
PP
avx
NP
q2
adj
PP
det
NP
q6
q7/1
npr
q8/1
prep
NP
PP
q9
q10/1

Figure 8

A Sample Recursive Transition Network

From (Woods 1970. P. 592).

Woods' next step along the dimensions of adding control was to generalize these recursive transition networks further and provide for both preference ordering on the arcs exiting a given node and arbitrary predicates which act as filters and which must be satisified for the node to be pursued. See Figure 2 for a piece of such an augmented transition net.

With the additional augmentation of allowing the system to maintain registers holding various pieces of information found during the parse, the ATN formalism has the full capabilities of a transformational grammar, but incorporates these abilities in a form that admits of far more efficient parsing algorithms than those previously proposed for transformational grammars. The addition of more powerful control information is only one of the arguments Woods advances for his ATN approach. Others relate to the perspicuity of such a representation, the efficiency involved in merging grammatical rules into common states, and the power obtained from the use of registers to allow context dependence and flexible structure-building activities. Thus, it is possible to view ATN's as basically addressed to the issue of control. The grammar is organized into states representing natural decision points. The possibilities exiting a state are ordered (as in production systems) and tests are placed on each arc to determine its admissibility. The resulting formalism has proved to be successful and now serves as the basis of Woods' lunar system (Woods 1974) and several systems constructed by Simmons (1973, 1975).

Winograd's SHRDLU also addressed the control issue. Woods added control knowledge to a state-oriented description. Winograd added control knowledge by pursuing an entirely procedural approach, representing grammatical knowledge as programs. These programs, being procedures, could have conditionals, loops, variables, or recursion. Figure 3 illustrates some simplified programs for recognizing sentences, noun phrases, and verb phrases (expressed as flowcharts rather than in Winograd's Programmar formalism).

Woods preferred to preserve some element of declarative representation by clearly naming the states; Winograd preferred the freedom of an entirely procedural description. The approaches are computationally equivalent (Winograd 1972) and both represent the fashion in which first-order grammatical theories can be extended to support efficient computation. Of course, most AI theories are computationally equivalent, just as most programing languages can, through more or less labor, be made to emulate each other. This is not, however, the real issue. The critical question is the perspicuity of the language or representation for the particular applications envisioned. Woods emphasizes the virtues of a net description; Winograd prefers expression directly as programs.

## An Issue for the Future--Deterministic Parsing

Delving into a particular theoretical issue helps to illustrate the nature of these ideas. A concern for control structure raises an obvious question for the future: Will it be possible to organize grammatical knowledge (including interactions with other kinds of knowledge) in such a way that backup is eliminated (without recourse to parallelism, which is ultimately equivalent to backup, at least for those sentences which, one intuitively feels, do not have a multitude of interpretations?

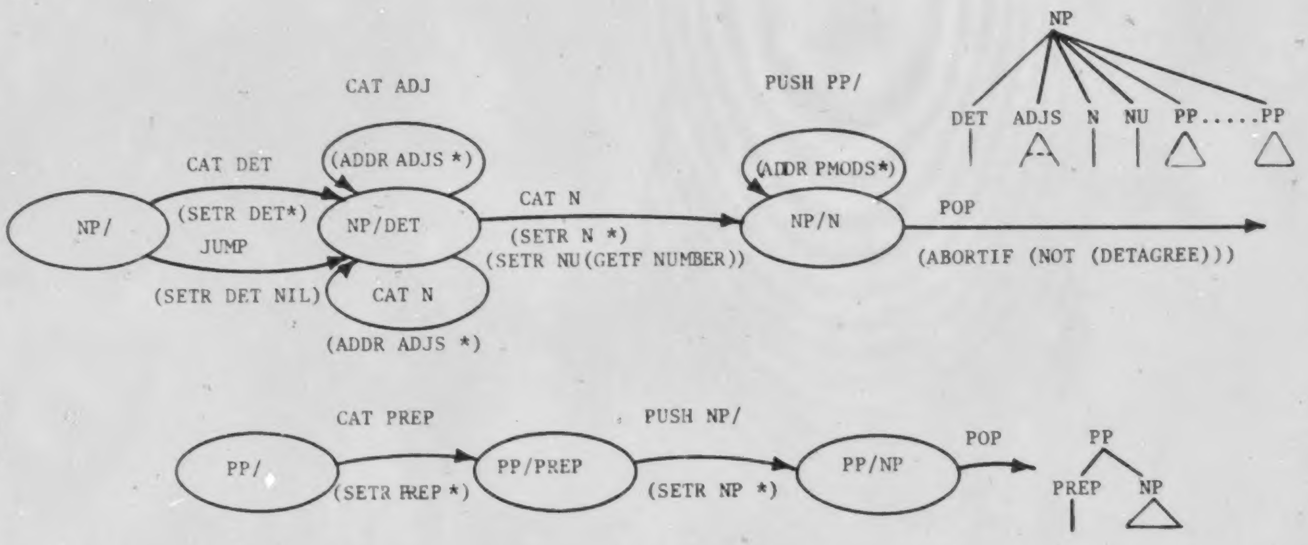

NP
DET
ADJS
N
NU
PP.....PP
CAT ADJ
PUSH PP/
CAT DET
(ADDR ADJS *)
(ADDR PMODS*)
CAT N
NP/
(SETR DET*)
JUMP
NP/DET
NP/N
POP
(SETR N *)
(SETR NU(GETF NUMBER))
(ABORTIF (NOT (DETAGREE)))
(SETR DET NIL)
CAT N
(ADDR ADJS *)
CAT PREP
PUSH NP/
POP
PP
PP/
PP/PREP
PP/NP
(SETR PREP *)
(SETR NP *)
PREP
NP

Figure 9

A Small Transition Network Grammar

From (Woods 1974. P. 137)

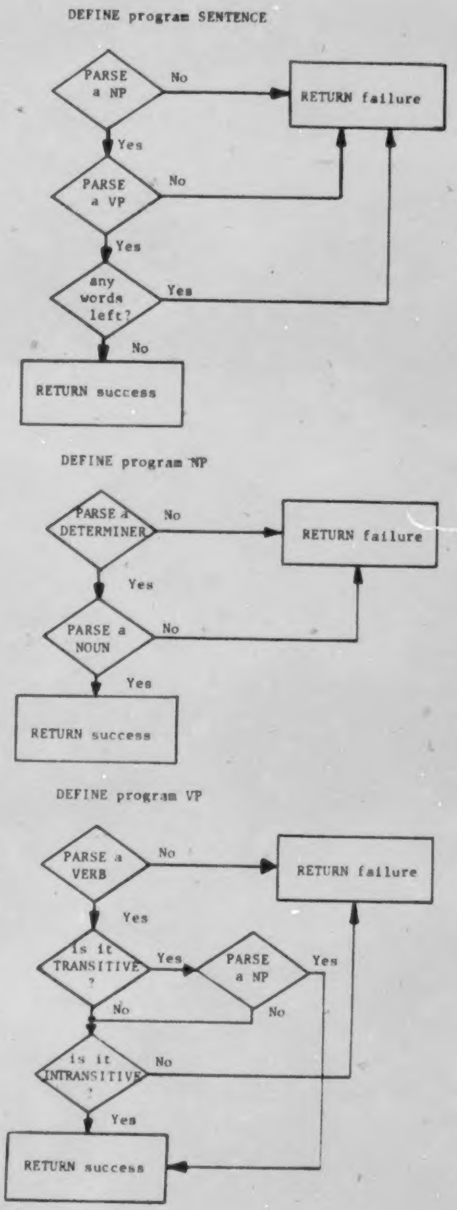

DEFINE program SENTENCE
PARSE a NP
No
RETURN failure
Yes
PARSE a VP
No
Yes
any words left?
Yes
No
RETURN success
DEFINE program NP
PARSE a DETERMINER
No
RETURN failure
Yes
PARSE a NOUN
No
Yes
RETURN success
DEFINE program VP
PARSE a VERB
No
RETURN failure
Yes
Is it TRANSITIVE?
Yes
PARSE a NP
Yes
No
No
Is it INTRANSITIVE?
No
Yes
RETURN success

Figure 10

Flowcharts of simplified PROGRAMMAR grammars for noun phrases, verb phrases, and sentences.

From (Winograd 1972. P. 83).

This goal of eliminating exhaustive search represents a natural concern of the computational epistemologist. First define the basic knowledge space. Then analyze its procedural structure. Viewing the process of finding a solution as a search branching at choice points, the question arises: "How is this branching controlled?" One possibility is to allow for backup. A second is to use parallelism. A third is to have diagnostic experts guide the decision at choice points so that backup is minimized or perhaps even eliminated.

This evolution has occurred most notably in the development of programs capable of symbolic integration. An early program by Slagle (1963) had knowledge of a basic set of integration formulas but combined this knowledge by means of a search. Recent programs following Moses' design of SIN (Moses 1967) have a deeper understanding of the nature of integration and do not require any search to find the integral. We do not necessarily expect language programs to evolve to the extreme form of deterministic problem solving that the integration programs have reached, but we do expect a similar kind of evolution in the same direction.

Insight into grammatical knowledge and knowledge representation techniques has reached the point where one can seriously consider a backup-free approach. The issue is not decided, but the suggestions are promising. For example, a backup-free scheme has recently been proposed by Marcus (1975). He calls his approach "wait and see parsing." The basic idea is that backup is unnecessary if the system is provided with a limited look-ahead of only a few words (or phrases) and if sufficient constraint is provided by semantic and pragmatic knowledge. He is currently constructing a case-frame-based parser following these guidelines.

Another proposal by Riesbeck (1975) involves a somewhat more radical approach. Riesbeck thinks that grammatical knowledge should be thoroughly integrated with semantic and pragmatic knowledge to such an extent that it is no longer reasonable to speak of separate components. This position represents the extreme along a natural line of evolution of more and more intimate interaction between grammatical and other kinds of knowledge. Riesbeck has constructed a comprehension system that reflects this philosophy. Words in the sentence trigger "productions" (called demons by Riesbeck) whose goal is to build a representation of the sentence's meaning (using Schank's conceptual dependency structures (Schank 1973)). Some of these productions reflect grammatical knowledge, such as generating the expectation for a noun following the appearance of an article. Others represent semantic and pragmatic information, such as expecting a human recipient following an appearance of the word "give." Riesbeck proposes that such an organization has both conceptual elegance as well as the ability to parse deterministically.

A new "conceptual" grammar by Martin (1975) also promises to generate perspicuous, backup-free conversions from sentences into expressions in his semantic representation. Martin's work suggests the possibility that an internal representation might be closer to the "surface" structure of natural language than most workers have believed feasible.

Clearly there is an important interaction between grammatical knowledge and other kinds of knowledge. A purely pipeline model in which a grammatical analysis is followed by a semantic analysis and, finally, a pragmatic analysis, is untenable. However, it is not yet clear which alternative among the following is preferable: (1) structuring the grammatical knowledge as a modular expert engaged in a dialog with other experts, or (2) distributing the grammatical knowledge with other kinds of knowledge in such a fashion that it is no longer appropriate to talk about separate syntactic, semantic, and pragmatic components. In any case, it is our expectation that in the next few years there will be language systems in which backup does not generally occur (except in the case of garden path sentences, which are equally confusing to people). This expectation is based on the observation that most procedural theories initially require backup, not because of the intrinsic properties of the problem area, but rather because sufficient knowledge to clarify alternatives has not yet been incorporated into the system. This knowledge has gradually been formalized (as our earlier discussion of work by Winograd and Woods showed). Consequently, we are evolving toward the stage where knowledge is available to make the right choices at each decision point.

## THE CLASSICAL, ROMANTIC, AND MODERN PERIODS OF AI

### A Change in Perspective

In the preceding chapters, we have tried to give the flavor of the new perspective from which AI approaches traditional problems in linguistics and epistemology. Specific forms of this general change in perspective are related to several theoretical issues:

- The relative importance of procedural, as opposed to "declarative" or "propositional" knowledge;
- The shift away from traditional logistical concepts (axioms, logical deduction) for the representation and manipulation of knowledge to more highly structured and particular representations;
- Liberation from the need for highly uniform transformational methods by allowing procedural sentries or "frame-keepers" to populate the knowledge system.

In each case we note a trend away from a traditional form toward a newer form. In each case experience has already shown us that such trends can go too far, and that there is often good reason for using the traditional forms. So, looking back over the brief history of AI, we see a dialectic process with three phases, which might be termed the Classical, Romantic, and Modern Periods of AI.

In the earliest work (which might be termed the Classical Period extending from the mid-fifties through the early sixties), knowledge is represented in very traditional ways: for example, as subject-predicate propositions or as important but relatively nonfundamental extensions of these propositions to property lists. Reasoning is accomplished through uniform techniques such as search or theorem proving.

Laymen often criticized typical programs of the Classical Period on the grounds that "people don't think like that." (Of course, this reaction may be totally wrong, since people do not necessarily know what deep processes lie under the conscious layers of thinking.) However, whether or not these objections are right in principle, the programs of what we shall term the Romantic Period (extending from the late sixties to the early seventies) do not evoke them nearly to the same extent. On the contrary, they sometimes appear to be nothing more than a new language for introspective material. Striking examples are the Evans ANALOGY program (Evans 1968) and the Winston learning program (Winston 1970).

This stage might be justifiably termed the Romantic Period for two reasons. One is the extent to which it represented an introspective endeavor, and the second is a notable tendency to argue for first one approach and then another as a complete theory of intelligence. The period began with an emphasis on networklike systems (Quillian 1967) and ended with a concentration on the representation of knowledge in an entirely procedural form (Hewitt 1972; Winograd 1972).

Finally, in this, the Modern Period, we begin to investigate more rigorously the conditions under which different kinds of knowledge are appropriate. Frames are extended from an essentially declarative data type to a procedural system through the incorporation of demons. Procedures, in order to allow debugging, are embedded in a web of declarative commentary specifying the purposes, requirements, bugs, and effects of programs. Procedural knowledge thus serves as a basis for action, while propositional knowledge provides a basis for understanding what one is doing and especially for knowing how to debug or change it.

## Some Unanswered Technical Issues in Frame Theory

In this section, we present a sampling of the questions that make the investigation of the ideas expressed in this essay into a scientific enterprise. The frame approach to representing knowledge is sufficiently novel that many of these questions are current issues of dispute in the community. We shall make no attempt to resolve them here. Our purpose is only to illustrate to the reader that AI is more than just a philosophic activity and that the use of computers requires that technical issues be resolved (and indeed serve to test the success of a given approach).

Learning. A theory that makes extensive use of prestored knowledge obviously faces a critical question: How is such knowledge obtained? Or, in other words, how does such a theory allow novel situations to be comprehended?

Reference and reasoning. Given that the system knows that feathers can be used to tickle (i.e., could successfully answer a question to that effect), it still remains to describe how indirect reference or reference through description is accomplished. Presumably, there are some tickling instruments that we have never before considered, but that we nevertheless are able to understand in that role. How are such inferences made?

Difficulties with sentries. A variety of technical problems arise in implementing the idea of a sentry as a demon. These include handling conflicts (i.e., situations where more than one demon is triggered by the current situation and each has an alternative meaning), forgetting (i.e., when should a demon become inactive), and context-switching (i.e., how are leaving and returning to contexts handled).

Representation. In many ways, the frame idea has served as a metaphor for us to present various ideas about the representation of knowledge, both declarative and procedural. However, many representation questions must be resolved in formalizing the notion. For example, how homogeneous is our representation of language to be? Are sentries (demons) to be represented in the same form as the basic frames? Are current computer languages such as LISP adequate, or do we require new programing constructs?

Control. Is knowledge of control implicit or explicit in a frame system? Control is explicit when, in addition to basic facts, the system also includes knowledge about when to apply these facts and how to resolve conflicts. Control thus becomes just another kind of knowledge to the system. On the other hand, control is implicit through the representation of knowledge as procedures, with ordering in the program, conditionals, and calls determining the sequence of application. Which approach is preferable?

Level interactions. We have posited the existence of frame-based knowledge for verbs, nouns, contexts, and themes. Technical questions arise about how these various levels of knowledge interact. Is the organization essentially hierarchical, with each level looking only at the results of the interpretation of the level below it? Or do all levels examine the text (perhaps in addition to examining each other)? Is communication occasionally heterarchical with higher levels sometimes directing the analysis of lower levels and at other times being in turn directed by the results of the lower level analysis?

This chapter represents the conclusion of our theoretical discussion of AI as a new style of epistemology and its impact on the understanding of language and comprehension. In the next chapter, we consider the relevance of these ideas to education.

# THE IMPACT OF ARTIFICIAL INTELLIGENCE ON EDUCATION

## If the Machines Get Smarter, Will the Children Get Smarter?

The relevance of artificial intelligence to education can be discussed along two different dimensions, corresponding to the narrow and broader views of the subject as "machine intelligence" and as "epistemology." The former view evokes the prospect of more useful machines, sometimes described under the label "intelligent computer-aided instruction" (Brown 1975). The other view has possible consequences of a much deeper sort, for it can be taken as projecting a total restructuring of the content and the theory of education. But even if this radical prediction is not fulfilled, the cognitive style of artificial intelligence has immediate and realistic suggestions about what children should be taught and how.

The two views can be stated in a more dramatic way. The organizers of the HUMRO-NSF conference on "Ten Year Predictions of the Impact of Computers on Education" (HUMRO 75) chose as the theme of one session: "If the machines get smarter, will the children get smarter?"

We think the deeper question is this: "Suppose that smart machines can be constructed. Should educators look to the machines themselves as instruments to make kids smart, or should they rather look at the principles which led to the machines becoming more intelligent to see whether these principles could be applied to children?" There is, of course, no need to choose. Our intention in so posing the question is to draw attention to the richer, but almost entirely neglected, second view. This paper is not the place for a full discussion of the relevance of artificial intelligence to education. Instead, we shall develop briefly three representative issues.

(1) The impact on education of the development of computer-based tutors with reasonable language comprehension abilities. We shall develop this topic under the heading "Intelligent Computer-Aided Instruction."

(2) The deeper question of whether the content of AI theories of language can have an impact on the very substance of linguistic studies in the schools, not just in the manner in which the current material is taught. We have chosen to call this topic "The White Box Approach to Language Studies."

(3) Whether the basic goal of AI in seeking unified theories of knowledge can be expected to have a profound impact on the entire educational curriculum, with the effect of decreasing the currently accepted division of education into disparate subject areas. We shall discuss this topic under the title "The Articulate Learner."

Again we remind the reader that our discussion is intended only to give the flavor of a few kinds of interaction that artificial intelligence and education might have. For a discussion of related areas of research in psychology, linguistics, and AI, see Winograd (1975). For a deeper discussion of the particular interactions discussed here, see Minsky-Papert (1974), Papert (1972), Goldstein (1974).

## Intelligent Computer-Aided Instruction (ICAI)

The question cited earlier, "If the machines get smarter, will the children get smarter?" immediately evokes the comment "certainly not automatically." This reaction follows from the deeper observation that the question is not primarily about machines, but rather about the nature of human intelligence and the effectiveness of teaching. To see this, transpose it into the two subquestions:

(1) "Could an idealized personal human tutor make a kid smarter?"

(2) "Could all or some of the functions performed by this tutor be mechanized?"

Answers to the first question will vary, depending on what one thinks about controversial issues in the theory of (human) intelligence. But there would certainly be less variation in answers to the question, "Could a highly intelligent, empathetic personal tutor enhance the intellectual development of children?" (Whether this enhancement is to be described as increasing the children's intelligence or the effectiveness of their use of it is not germane here.)

There are very few programs that can lay claim to fulfilling the functions of an intelligent, insightful tutor. However, important groundbreaking research has been done by Carbonell (1970) and Collins (1975) working on the SCHOLAR Project, and by Brown (1974) in his development of the SOPHIE program. The SCHOLAR Project has developed a series of successively more sophisticated computer-based tutors for such domains as geography. The project has involved the development of natural language comprehension modules, the construction of domain-oriented problem solvers based on a semantic net approach to representation, and, most notably, the procedural representation of tutorial strategies.

Recent work by Collins (1975) has involved a deep interaction between psychology and artificial intelligence in the form of studying the kinds of reasoning that people typically use to handle questions in open domains, i.e., concerning topics about which they know a great deal, but not all, of the relevant facts. People are typically quite good at making plausible inferences in such contexts. Collins' goal is to represent such reasoning strategies in a procedural way and then make that knowledge available to the SCHOLAR tutor for the purpose of allowing the tutor to help a student debug his own reasoning style.

Brown (1974) has developed a computer-based laboratory in which a student can learn electronic troubleshooting. The program SOPHIE contains an electronics problem solver that embodies a genuine understanding of the circuits involved, a natural language component that allows English question-answering to occur between the student and program, and an understanding of certain kinds of tutoring strategies. We shall not develop the details of Brown's program further, except to remark on the following points:

(1) The program is a milestone in CAI simply because of the extent to which it understands its subject matter. Previous CAI programs (and indeed most current ones) are notorious for their ignorance. The slightest deviation from the expected response on the part of the student in his answer results in the program erroneously judging the answer as incorrect. Brown's program employs a variety of AI techniques to simulate the circuits under study and predict accurately the consequences of various faults. This kind of domain competence is essential as a foundation for any CAI tutor, and we expect the incorporation of AI problem solvers into CAI programs to become more and more common in the coming years.

(2) The tutoring strategies of SOPHIE are concerned with the nature of reasoning involved in building up a model of a problem from incomplete evidence. This is a fundamental kind of human reasoning. It addresses the basic issue of what kinds of questions a student *should* ask in attempting to understand a new subject, device, or electronic circuit. SOPHIE is able to judge whether a measurement requested by the student (as part of the process of tracking down a fault in the circuit) is a reasonable question on the basis of what the student already knows. If not, i.e., if the student already possesses the information to answer the question, then the program draws the student's attention to the redundancy and discusses the information which the student knew but failed to consider.

This study of human cognitive processes through the formalisms developed by artificial intelligence and the incorporation of that understanding into CAI programs is still in the formative stages. But it is our hope that in the coming years, the concept of personal computer-based tutors available to any student at any time will become a reality.

## The White Box Approach to Language Studies

In the previous section, we saw the application of artificial intelligence research on language comprehension and problem solving to the design of intelligent CAI programs. In this section, we develop a more radical application.

The ICAI tutors, no matter how intelligent, are essentially *black boxes* from the student's perspective. The possibility exists that the concepts developed to design those programs might themselves be worthy objects of study by the students. We propose calling this a *white box* approach. Unlike the "black box" approach, the point here is precisely to encourage the young students to look inside the box, to ask how it works, and to gain insight into their own language by doing so. An example of the white box approach is illustrated by suggesting a new kind of language laboratory. We envision a computer-based environment in which students would have the opportunity to explore the structure of language by designing, programing, and testing various kinds of language comprehension programs.

The virtues of such a laboratory approach lie in the active nature of students' involvement, the possibility of students developing their ideas in a personal way not rigidly limited by the teacher's approach (but reasonably circumscribed by the goals of designing a successful program), and the exposure to new computational theories of language superior in many ways to traditional grammars as conceptual frameworks for the study of language at school.

As an example of the power of these new approaches to language, we mention the possibility that they hold out for a unified approach to the study of knowledge. This possibility was illustrated earlier by our discussion of frame theory in which the same basic knowledge-representation techniques applied to grammatical knowledge as well as other kinds of knowledge concerned with themes, contexts, and semantic meanings. The exciting possibility exists of diminishing the compartmentalization of school knowledge. Linguistic knowledge becomes less of a "different" thing, apart from other knowledge.

Our series of roles for language understanding machines has moved progressively away from a focus on the machine to a focus on the linguistic idea that make the machine possible. We conclude this essay by generalizing this approach beyond the linguistic domain.

## The Articulate Learner

In this section, we develop further the notion of a unified approach to educational curricula by introducing the notion of articulate learning. Our discussion is motivated by the question why students who find language studies in school challenging and who do well in them often find the mathematics curriculum a confusing and impossible maze. A traditional answer is that there is such a thing as quantitative mathematical abilities and qualitative language-oriented abilities. Such a belief is mirrored in all of the achievement tests that students are asked to take in which there are clearly demarcated language and mathematics sections.

We would like to raise the possibility that the traditional mathematics curriculum of grades 1-12 lacks certain dimensions which make it understandably obscure compared to language studies. In particular, its basic failing is in the extreme difficulty it presents for a student or teacher who attempts to be articulate about the problem-solving process.

In language studies, a teacher can discuss the many dimensions of a story written by a student: its grammar, its theme, its narrative structure, the plot, the setting. In contrast, what can a teacher discuss about the failure of a student to solve an arithmetic problem or an integral? For example, we have seen a child go to the blackboard and solve an arithmetic problem in the following way:

$$\begin{array}{r} 35 \\ +\ 35 \\ \hline 610 \end{array}$$

The teacher involved had little to say besides observing that the answer was wrong. We believe that the earlier discussion of AI and epistemology has some very definite suggestions for altering the mathematics curriculums to allow the student and the teacher to be far more articulate about the knowledge and problem solving involved.

For the above example, we imagine an environment in which the first kind of knowledge to which a student is exposed is the structure of programs and the related ideas of planning and debugging. Skill in designing procedures lies at the heart of achieving competence in mathematics (and in writing essays as well). Given access to a vocabulary for programs, plans, and bugs, we believe that student and teacher could be articulate about describing the particular algorithm used by the student in reaching the "610" answer, in identifying the bugs, and in debugging the addition program to yield a correct result. For the above example, we imagine a discussion similar to the earlier description of the HACKER program's reasoning regarding its problem solving in the Blocks World. The student had mastered an addition procedure for single digits. His goal in this problem is to apply that procedure as a subroutine to the more complex goal of adding multidigit numbers. His solution indicates that he followed the reasonable linear plan of assuming that the sum of multidigit numbers could be achieved by adding the columns independently. Linear planning commits the problem solver to have some method of putting together the solutions of the subproblems at which he arrived independently. The answer, "610," indicates that our student used the simple artifice of concatenating the two partial solutions. The bug, of course, was in assuming the complete independence of the column additions. The debugging is accomplished by taking account of this interaction through ordering the columnar additions right-to-left and employing a "carry" data structure.

The reader may find a vocabulary involving such terms as linearity, bug, debugging, independence, and data structure confusing; and must certainly believe that such a vocabulary is hardly usable with an elementary school student. This is true, of course, given the current curriculum. However, our proposal is not simply to add this vocabulary to the traditional curriculum. The proposal is to redesign the curriculum to allow introduction of this and related concepts at an early age.

This is being done in a preliminary way at MIT's LOGO project (Papert 1972) in which new computation-based approaches to mathematics, science, and music are currently being developed. As an example, elementary school students are introduced to the above set of programing concepts and to traditional mathematical concepts in the context of programing a computer display to draw pictures of their choice (Figure 11).

The children construct the programs themselves in LOGO, a programing language expressly designed for students. A successful LOGO experience involves not only self-motivation of the student to construct such programs (because of the enjoyment of creating pictures), but also involves discussions between student and teacher, between students and students, and among the students themselves, in which students become verbal about their reasoning and learn to use a vocabulary describing their plans, their programs, and, most important, their bugs.

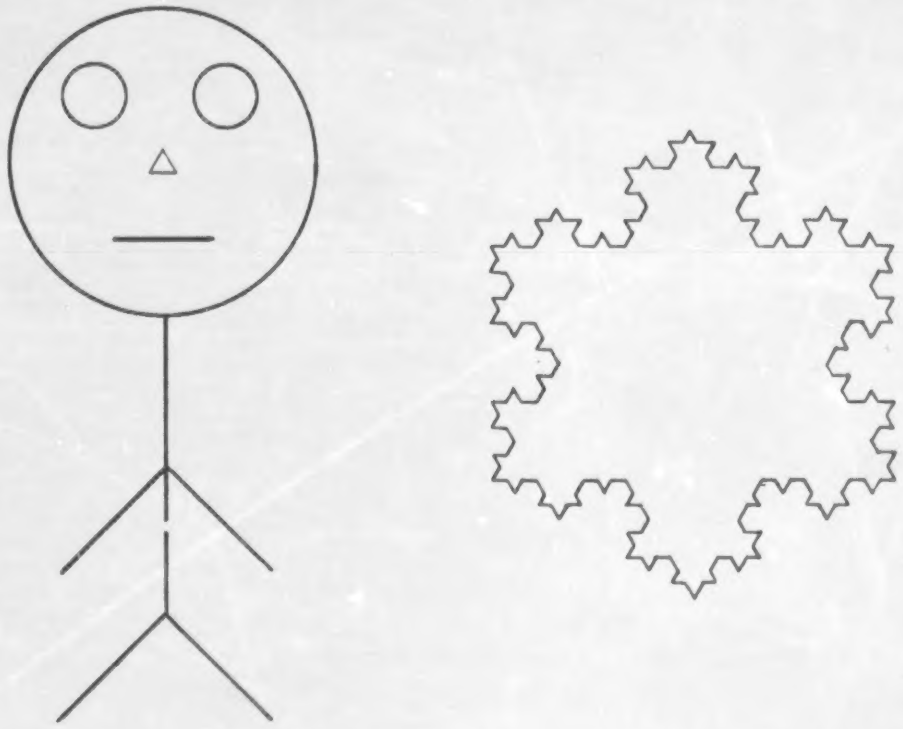


Figure 11

Typical Pictures Drawn by Children

From (Goldstein 1974. Pp. 9 and 21).

The design of new educational environments in which the student is introduced to the concepts that AI has found essential to the construction of intelligent programs, is only in a beginning state. But our experience with children in teaching them Turtle Geometry, computer-based physics, and computer-based composing leads us to believe that the approach is both revolutionary and promising. Furthermore, it seems to lead to a situation where students can be articulate about their own problem solving, regardless of whether it is in the mathematical or linguistic domain.

## Epistemology and Education

The role of a knowledge-oriented AI approach to the study of language comprehension has its counterparts in all areas of cognitive science, and each of these carries implications for education.

We have argued again and again that a major thrust of AI research has been to find representations of knowledge that lend themselves to effective integration into larger knowledge systems. The analogous problem for education is obvious: Are the representations of knowledge that have proven themselves effective in augmenting machine intelligence also of value in augmenting human intelligence? And, if not, does a similar inquiry into the forms of knowledge suitable for human thinking yield a different set of modes of representation? Furthermore, the epistemological paradigm for the study of learning asks questions like: What knowledge does a good learner use in a given situation? What knowledge would make learning a simpler process than it appears to an observer who does not know this knowledge is being used, or to a learner who lacks it?

Very little attention has been paid to problems of this sort. But signs of a rich area of the utmost importance seem to be present in the results of such preliminary investigations as the construction of ICAI programs containing procedural models of human inference and learning, the design of new approaches to learning based on debugging theory, and the design of new curriculums such as Turtle Geometry, based on the idea of the articulate learner.

It is our belief that future explorations into the nature of language, of various subject domains, and of the elements of intelligence itself--if pursued from an epistemological perspective grounded in the computational metaphor--will provide the foundation for deeper theories of the nature of language and of the nature of learning. Furthermore, we expect such theories to have practical consequences in the form of programs that comprehend language and in the design of new kinds of computer-based educational environments.

# EPILOG

## Controversies About Educability and Innateness

The general issues about the theory of education discussed in the last chapter cut across a number of acute current controversies about the relations between linguistic abilities in the developing child and other cognitive functions including learning. Different approaches to linguistics and to cognitive theory appear to lead theorists to very different opinions on questions such as the domain-specificity of language skills; their relation to particular brain structures; and the extent of their innateness. For example, one sees a very sharp opposition between the Chomskian and Piagetian lines of thinking (Chomsky 1975). For the former, there is a definite linguistic competence extremely specific to language (i.e., has very little overlap with other cognitive abilities) and determined by innate biological structures. Adherents of the Piagetian school, on the other hand, tend to see linguistic competence as sharing major cognitive processes with other intellectual domains and to be explicable more in terms of a psycho-developmental process. Any such brief summary of complex theories necessarily oversimplifies the respective positions. Nevertheless, the oppositions, though more modulated, are real and of fundamental and urgent importance.

We see work in artificial intelligence as having a direct bearing on this issue. We do not mean that we are able to resolve the issues. Rather we believe that computational models provide a theoretical overview necessary even for a clear statement of the issues. This belief is reinforced by the increasingly rapid spread of computational ideas into the work of psychologists and even philosophers whose views fall on both sides of these controversies. See Fodor (1975) for an example of an extreme "nativist" position formulated within a computational theoretical frame by a psychologist who disagrees fundamentally with much of the "mainline" AI trends. See Norman (1975) and Bobrow (1975) for views of psychologists sympathetic to the kind of AI theory treated in this paper.

In conclusion, though the debates about intelligence and cognition rage around AI, and the reader may have objected to many of the specific observations that we have made regarding frame theory and other issues, we believe that there already has been this fundamental and profound change: That the language of computation has become the proper dialect for discussing the basic issues of both psychology and education.

## Bibliography

Anderson, J. R. and Bower, G. Human Associative Memory. Washington, D. C.: H. V. Winston, 1973.

Brown, J. S., Burton, R. R., and Bell, A. G. SOPHIE: A Sophisticated Instructional Environment for Teaching Electronic Troubleshooting. BBN Report No. 2790. Cambridge: Bolt, Beranek, and Newman, 1974.

Brown, J. S. Proposal for Knowledge-based Intelligent Computer Assisted Instruction. BBN Proposal No. P75-CSC-17A. May 28, 1975. Cambridge: Bolt, Beranek, and Newman, 1975.

Bobrow, D. and Collins, A. (eds.). Representation and Understanding. New York: Academic Press, 1975.

Carbonell, H. AI in CAI: An Artificial Intelligence Approach to Computer-Assisted Instruction. IEEE Trans. Man-Machine Systems, MMS-11, 190-202, December 1970.

Celce-Murcia, M. English Comparatives. Unpublished doctoral dissertation, Dept. of Linguistics, UCLA, Los Angeles, 1972.

Chafe, W. L. Meaning and the Structure of Language. Chicago: Univ. of Chicago Press, 1970.

Charniak, E. Toward a Model of Children's Story Comprehension, AITR-266. Artificial Intelligence Laboratory, Cambridge, MIT, 1972.

Chomsky, N. Syntactic Structures. The Hague: Uitgeverij Mouton, 1957.

Chomsky, N. Questions of Form and Interpretation. Linguistic Analysis, 1, 1, 1975.

Collins, A., Warnock, E. H., and Passafiume, J. J. Analysis and Synthesis of Tutorial Dialogues. In G. Bower (ed.). The Psychology of Learning and Motivation, Vol. 9, New York: Academic Press, 1975.

Collins, A., Warnock, E. H., Aiello, N., and Miller, M. Reasoning from Incomplete Knowledge. In D. Bobrow and A. Collins (eds.). Representation and Understanding. New York: Academic Press, 1975.

Drake, C. Semantic Case Frames in the Understanding of Connected Discourse. Unpublished masters dissertation. Cambridge, MIT, 1975.

Evans, T. A Program for the Solution of Geometric-Analogy Intelligence Test Questions. In M. Minsky (ed.). Semantic Information Processing. Cambridge: MIT Press, 1968.

Fillmore, C. The Case for Case. In E. Bach and R. T. Harms (eds.). Universals in Linguistic Theory. New York: Holt, Rinehart, and Winston, 1968.

Fodor, J. A. The Language of Thought. New York: T. Y. Crowell, 1975.

Goldstein, I. P. Understanding Simple Picture Programs, AITR-294. Artificial Intelligence Laboratory, Cambridge, MIT, 1974.

Green, B. F., Wolf, A. K., Chomsky, C., and Laughery, K. Baseball: An Automatic Question Answerer. In E. A. Feigenbaum and J. Feldman (eds.). Computers and Thought. New York: McGraw-Hill, 1963.

Heidorn, G. Augmented Phrase Structure Grammar. In R. Schank and B. L. Nash-Webber (eds.). Proceedings of a Workshop on Theoretical Issues in Natural Language Processing. Cambridge, MIT, 1975.

Hewitt, C. Description and Theoretical Analysis (Using Schemata) of Planner, AITR-258. Artificial Intelligence Laboratory, Cambridge, MIT, 1972.

Human Resources Research Organization. Ten Year Forecast for Computers and Communications: Implications for Education. Warrenton, Va.: Arlie House, September 1975.

Katz, J. J. and Fodor, J. A. The Structure of a Semantic Theory. In J. A. Fodor and J. J. Katz (eds.). The Structure of Language. Englewood Cliffs: Prentice-Hall, 1964.

Lindsay, R. K. Inferential Memory as the Basis of Machines which Understand Natural Language. In E. A. Feigenbaum and J. Feldman (eds.). Computers and Thought. New York: McGraw-Hill, 1963.

Marcus, M. Diagnosis as a Notion of Grammar. In R. Schank and B. L. Nash-Webber (eds.). Proceedings of a Workshop on Theoretical Issues in Natural Language Processing. Cambridge, MIT, 1975.

Martin, W. Conceptual Grammar, Internal Memo 20. Automatic Programming Group, Project MAC. Cambridge, MIT, 1975.

Minsky, M. and Papert, S. Artificial Intelligence. Condon Lectures, Oregon State System of Higher Education, Eugene, Oregon, 1974.

Minsky, M. A Framework for Representing Knowledge. In P. Winston (ed.). The Psychology of Computer Vision. New York: McGraw-Hill, 1975.

Moses, J. Symbolic Integration. MAC-TR-47, Project MAC. Cambridge, MIT, 1967.

Newell, A. Artificial Intelligence and the Concept of Mind. In R. C. Schank and K. M. Colby (eds.). Computer Models of Thought and Language. San Francisco: Freeman, 1973.

Norman, D. A. and Rumelhart, D. E., and the LNR Research Group. Explorations in Cognition. San Francisco: Freeman, 1975.

Papert, S. Teaching Children Thinking. Programmed Learning and Educational Technology. 9, 5, September 1972.

Quillian, M. R. Semantic Memory. In M. Minsky (ed.). *Semantic Information Processing*. Cambridge: MIT Press, 1968.

Riesbeck, C. K. Computational Understanding: Analysis of Sentences and Context. Working Paper, *Fondazione Dalle Molle per gli studi linguistici e di communicazione internazionale*. Castagnola, Switzerland, 1974.

Riesbeck, C. K. Computational Understanding. In R. Schank and B. L. Nash-Webber (eds.). *Proceedings of a Workshop on Theoretical Issues in Natural Language Processing*. Cambridge, MIT, 1975.

Rumelhart, D. E. Notes on a Schema for Stories. In D. Bobrow and A. Collins (eds.). *Representation and Understanding*. New York: Academic Press, 1975.

Schank, R., Goldman, N., Rieger, C., and Riesbeck, C. Primitive Concepts Underlying Verbs of Thought. AIM-162. Stanford: Stanford Univ., 1972.

Schank, R. C. Identification of Conceptualizations Underlying Natural Language. In R. C. Schank and K. M. Colby (eds.). *Computer Models of Thought and Language*. San Francisco: Freeman, 1973.

Schank, R. C. Using Knowledge to Understand. In R. Schank and B. L. Nash-Webber (eds.). *Proceedings of a Workshop on Theoretical Issues in Natural Language Processing*. Cambridge, MIT, 1975.

Simmons, R. F. Semantic Networks: Their Computation and Use for Understanding English Sentences. In R. C. Schank and K. M. Colby (eds.). *Computer Models of Thought and Language*. San Francisco: Freeman, 1973.

Simmons, R. F. The Clown's World. In R. Schank and B. L. Nash-Webber (eds.). *Proceedings of a Workshop on Theoretical Issues in Natural Language Processing*. Cambridge, MIT, 1975.

Slagle, J. R. A Heuristic Program that Solves Symbolic Integration Problems in Freshman Calculus. In E. A. Feigenbaum and J. Feldman (eds.). *Computers and Thought*. New York: McGraw-Hill, 1963.

Sussman, G. J. *A Computational Model of Skill Acquisition*. New York: American Elsevier, 1975.

Wilks, Y. Natural Language Understanding Systems within the A. I. Paradigm: A Survey and Some Comparisons. Unpublished paper prepared for the National Institute of Education.

Winograd, T. *Understanding Natural Language*. New York: Academic Press, 1972.

Winograd, T. Frame Representations and the Declarative/Procedural Controversy. In D. Bobrow and A. Collins (eds.). *Representation and Understanding*. New York: Academic Press, 1975.

Winograd, T. Artificial Intelligence and Language Comprehension. Unpublished paper prepared for the National Institute of Education, 1975.

Winston, P. Learning Structural Descriptions by Examples. Reprinted in P. Winston (ed.). The Psychology of Computer Vision. New York: McGraw-Hill, 1975.

Woods, W. Transition Network Grammars for Natural Language Analysis. Communications of the ACM, 13, 10, October 1970.

Woods, W. Natural Language Communication with Computers, BBN Report No. 2976. Vol. 1. Cambridge: Bolt, Beranek, and Newman, 1974.

# NATURAL LANGUAGE UNDERSTANDING SYSTEMS WITHIN THE AI PARADIGM: A SURVEY AND SOME COMPARISONS

Yorick Wilks
Department of Artificial Intelligence
University of Edinburgh
Edinburgh, Scotland

This paper surveys the major projects on understanding natural language that fall within what may now be called the artificial intelligence paradigm of natural language systems. Some space is devoted to arguing that the paradigm is now a reality and different in significant respects from the generative paradigm of present day linguistics. The comparisons among systems center around questions about the level, centrality, and "phenomenological plausibility" of the knowledge and inferences that must be available to a system that is to understand everyday language.

The research and writing of this paper were sponsored under NIE Order No. NIE-P-75-0026. The views and conclusions contained are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of the National Institute of Education or the United States Government.

# TABLE OF CONTENTS

INTRODUCTION 81

WINOGRAD'S UNDERSTANDING SYSTEM 83

SOME DISCUSSION OF SHRDLU 85

SOME MORE GENERAL BACKGROUND ISSUES 87

SECOND GENERATION SYSTEMS 89

SOME COMPARISONS AND CONTRASTS 104

CONCLUSION 113

REFERENCES 115

# Introduction

In his report to the Science Research Council on the state of artificial intelligence, Sir James Lighthill (1973) gave most of the field a rather bad prognosis. One of the few hopeful signs he saw was Winograd's (1972) natural language understanding system. Yet now, only a year later, Winograd has stopped work on the system he constructed, and has begun a new system, based on entirely different principles (see Winograd 1974). He went so far, in a survey lecture (Winograd 1973) of extraordinary modesty (in a field not known for its small claims), to place his celebrated early work in only the first generation of computer systems designed to understand natural language, and went on to describe others' second generation systems.

I shall return later to this metaphor of generations, but what is one to say in general terms of a field where yesterday's brightest spots are today's first generation systems, even though they have not been criticized in print, nor shown in any generally acceptable way to be fundamentally wrong? Part of the answer lies in the profound role of fashion in artificial intelligence in its present prescientific phase. A cynical American professor remarked recently that artificial intelligence (AI) had an affair with someone's work every year or two, and that, just as there were no *reasons* for falling in love, so, later, there were no reasons for falling out again. In the case of Winograd's work it is important now to resist this fashion, and reemphasize what a good piece of research it was, as I shall in a moment.

Another part of the answer lies in the still fundamental role of *metaphysical criticism* in AI. In the field of computer vision things are bad enough, in that anybody who can see feels entitled to criticize a system, on the ground that he is sure he does not see using such and such principles. In the field of natural language understanding things are worse: Not only does anyone who can speak and write feel free to criticize on the corresponding grounds, but in addition there are those trained in disciplines parasitic upon natural language, linguists and logicians, who often know in addition how things MUST BE DONE on a priori grounds. It is this metaphysical aspect of the subject that gives its disputes their characteristically acrimonious flavor.

In this paper I want to sort out a little what is agreed and what is not; what are some of the outstanding disputes and how testable are the claims being made. If what follows seems unduly philosophical, one should remember that little *is* agreed, and almost no achievements are beyond question. To pretend otherwise, by concentrating only on the details of established programs, would be meretricious and misleading.

To survey an energetic field like this one is inevitably to leave a great deal of excellent work unexamined, at least if one is going to do more than give a paragraph to each research project. I have left out of consideration at least six groups of projects:

(1) Early work in artificial intelligence and natural language that has been surveyed by Winograd (1973) and Simmons (1970) among others.

(2) Work by graduate students of, or intellectually dependend upon that of, people discussed in some detail here.

(3) Work that derives essentially from projects described in detail here. This group embraces several groups interested in testing psychological hypotheses, as well as others constructing large-scale systems for speech recognition. I have devoted no space to speech recognition as such here, for it seems to me to depend upon the quality of semantic and inferential understanding as much as anything, and so I have concentrated upon this more fundamental task.

(4) Work on language generators, as opposed to analyzers and understanders. They are essential for obtaining any testable output, but are theoretically secondary.

(5) All the many and varied reasoning schemes now available in AI, including PLANNER (Hewitt 1969), QA4 (Rulifson et al. 1972), MERLIN (Moore and Newell 1973), as well as automatic programing (Balzer et al. 1974) (Heidorn 1974), and debugging (Sussman 1974) projects, many of which are producing formalisms that appear increasingly like natural language.

(6) Conservative reasoning schemes, such as first order predicate calculus, that have been applied to, or advocated for, the analysis of natural language, such as McCarthy and Hayes (1969), Coles (1972), and Sandewall (1972).

The exclusions under (2) above are particularly unfair to the work in the unpublished theses of Rieger (1974) and McDermott (1974), and can only be justified, like those of sections (5) and (6) above, by space, bias, and the considerations advanced in a later section of this survey concerning what it is for AI research to be about natural language, rather than about something else like human psychology, or logic, or the bargaining behavior of children under stress.

One final remark about the style of this survey. I have not indulged the normal critical luxury of careful and extensive quotation from authors followed by detailed and meticulous criticism. That would be wholly inappropriate in the present state of things. A great deal of work is being done at the moment, and many of the principal researchers change their views on very fundamental questions between one paper and the next without drawing any attention to the fact. Cheap self-contradictions and changes of mind are all too easy to find, so criticism and comparisons are best drawn with a very broad brush and a light stroke.

## Winograd's Understanding System

Winograd's system runs as a dialog, in real time, between a human operator who types in messages, and the system proper, called SHRDLU, which displays on a screen pictures of a closed world of colored blocks and pyramids, a box in which to put them, and an arm or crane that can move the objects about. SHRDLU displays its replies to the operator by writing on the same screen. The following is a discontinuous sample of the dialog Winograd gives in his book. SHRDLU's replies are in upper case:

> pick up the big red block
>
> OK
>
> find a block which is taller than the one you are holding and put it into the box
>
> BY 'IT,' I ASSUME YOU MEAN THE BLOCK WHICH IS TALLER THAN THE ONE I AM HOLDING
>
> OK
>
> what does the box contain
>
> THE BLUE PYRAMID AND THE BLUE BLOCK
>
> had you touched any pyramid before you put the green one on the little cube
>
> YES, THE GREEN ONE
>
> when did you pick it up
>
> WHILE I WAS STACKING UP THE RED CUBE, A LARGE RED BLOCK, AND A LARGE GREEN CUBE
>
> why
>
> TO GET RID OF IT.

This ingenious program was also able to accept new definitions of words, such as "steeple," to make use of what it knew about the properties of the objects described, such as that a block could have its top cleared off, but a pyramid could not; and also to remember what it had done before, as in the sample above.

The program was written in the language PLANNER (actually, in a subset of PLANNER called PROGRAMMAR), which is a concrete expression of the slogan "meanings are procedures," a sentiment into whose own meaning it is probably best not to inquire too closely, but which has undoubtedly led to a new style of programing. PLANNER is a theorem-proving language: It tries to establish the truth of assertions, not in the normal uniform, proof-theoretic manner, but by accepting a range of "programed hints" about how best to proceed at any point. In a language understanding program like Winograd's, this means replacing familiar grammar rules such as S → NP + VP (a sentence consists of a noun phrase followed by a verb phrase) by procedures, in this case: ((PDEFINE SENTENCE ((PARSE NP) NIL FAIL) ((PARSE VP) FAIL FAIL RETURN))).

The details of the notation need not detain us; what is important is that Winograd's grammar is not the conventional list of rules, but small subprograms like the lines above, that actually represent *procedures for imposing* the desired grammatical structure. The definitions of more complex words are also in this form: here, for example, is the theorem defining the content of "pickup":

```
(DEFTHEOREM TC-PICKUP (THCONSE (X(WHY (EV))EV)

(ᵘPICKUP $?X) (MEMORY)(THGOAL(ᵘGRASP $?X) (THUSE TC-GRASP))
(THGOAL (ᵘRAISEHAND)(THNODB)(THUSE TC-RAISEHAND))
(MEMOREND(ᵘPICKUP $?EV  $?X))))
```

Once again the details of the notation need not be explained in order to see that the word is being defined in terms of a number of more primitive subactions, such as RAISEHAND, each of which must be carried out in order that something may indeed be picked up. The linguistic content is a vocabulary of what seems to be about 175 words, a "systemic grammar," due to M. A. K. Halliday (1970), plus a simple system of semantic "features," marking words and arranged hierarchically, such as PHYSOB (for physical object words) and ANIMATE (for animate words like "robot") together with some factual knowledge about the block world. Both types of knowledge, linguistic and factual, are represented in PLANNER. When the program actually runs, it is able to access whichever sort of knowledge is required at any given moment, rather than in the conventional manner of first doing syntactic parsing to get a syntactic structure and then manipulating the features to get a semantic structure.

One reason for the enormous impact of this work was that, prior to its appearance, AI work was linguistically trivial, while the systems of the linguists had no place for the use of inference and real world knowledge. Thus a very limited union between the two techniques was able to breed considerable results. Before Winograd there were few programs in AI that could take a reasonably complex English sentence and ascribe any structure whatever to it. In early classics of natural language understanding in AI, such as Bobrow's STUDENT (1968) problem solver for simple algebra, input sentences had to be short and of stereotyped form, such as "what is the sum of . . .?"

Conversely, in linguistics there was, until very recently, little speculation on how we understand the reference of pronouns in such elementary sentences as "the soldiers fired at the women and I saw *several* fall," where it is clear both that the answer is definite, and that finding it requires some inferential manipulation of generalizations about the world. Readers should ask themselves at this point *how* they know the referent of the pronoun in that sentence.

## Some Discussion of SHRDLU

So far, the reaction to Winograd's work has been wholly uncritical. What would critics find to attack if they were so minded? First, that Winograd's linguistic system is highly conservative, and that the distinction between "syntax" and "semantics" may not be necessary at all. Secondly, that his semantics is tied to the simple referential world of the blocks in a way that would make it inextensible to any general, real world, situation. Suppose "block" were allowed to mean "an obstruction" and "a mental inhibition," as well as "a cubic object." It is doubtful whether Winograd's features and rules could express the ambiguity, and, more important, whether the simple structures he manipulated could decide correctly among the various meanings in any given context of use. Again, far more sophisticated and systematic case structures than those he used might be needed to resolve the ambiguity of "in" in "He ran the mile in five minutes," and "He ran the mile in a paper bag," as well as the combination of case with word sense ambiguity in "He put the key in the lock" (door lock) and "He threw the key in the lock" (river lock).

The blocks world is also strongly deductive and logically closed. If gravity were introduced into it, then anything supported that was pushed in a certain way would logically *have* to fall. But the commonsense world of ordinary language, is not like that: In the "women and soldiers" example given earlier, the prounoun "several" can be said to be resolved using some generalization such as "things shot at and hurt tend to fall." There are no logical "have to's" there, even though the meaning of the pronoun is perfectly definite.

Indeed, it might be argued that, in a sense, and as regards its semantics, Winograd's system is not *about* natural language at all, but about the other technical question of how goals and subgoals are to be organized in a problem-solving system capable of manipulating simple physical objects. If one glances back at the definition of "pickup" quoted above, one can see that it is in fact an expression of a procedure for picking up an object in the SHRDLU system. Nothing about it, for example, would help one understand the perfectly ordinary sentence "I picked up my bags from the platform and ran for the train." One could put the point so: What we are given in the PLANNER code is not a *sense* of "pick up" but a case of its use, just as "John picked up the volunteer from the audience by leaning over the edge of the stage and drawing her up by means of a rope clenched in his teeth" is not so much a sense of the verb as a use of it.

Those who like very general analogies may have noticed that Wittgenstein (1953, para. 2ff.) devoted considerable space to the construction of an elementary language of blocks, beams, and slabs; one postulated on the assumption that the words of language were basically, as is supposed in model theory, the names of items. But, as he showed of the enterprise, and to the satisfaction of many readers, "That philosophical concept of meaning (i.e., of words as the unambiguous names of physical objects--YW) has its place in a primitive idea of the way language functions. But one can also say that it is the idea of a *language more primitive than ours*." (My italics)

To all this, one might counter that it has not been shown that the language facilities I have described *cannot* be incorporated in the structures that SHRDLU manipulates, and that, even if they could not, the work would still be significant in virtue of its original control structure and its demonstration that real world knowledge can be merged with linguistic knowledge in a working whole. Indeed, although Winograd has not tried, in any straightforward sense, to extend the SHRDLU system, one could say that an extension of this sort is being attempted by Brown (1974) with his "Believer System," a hybrid system combining a component about beliefs that is second generation (see page 89), with a base analyzer from Bruce's CHRONOS system (1972) which is a microworld--late first generation--system in the same sense as Winograd's. Others in the last category that should be mentioned are Davies and Isard's (1972) exploration of the concepts of "must" and "could" in a microworld of tic-tac-toe, and Joshi's extension of it (1973), but above all the important and influential work of Woods (1972).

This work, most recently applied to a microworld of lunar rock samples, is not discussed in the detail it deserves in this paper. The system, based on an augmented state transition network grammar, is undoubtedly one of the most robust in actual use, in that it is less sensitive to the *particular* input questions it encounters than its rivals. The reason for not treating it in depth is that both Woods and Winograd have argued in print that their two systems are *essentially equivalent* (Winograd 1971) (Woods 1973); if they are right, there is no need to discuss both, and Winograd's is, within the AI community at least, the better known of the two.

Their equivalence arguments are probably correct: Both are grammar-based deductive systems, operating within a question-answering environment in a highly limited domain of discourse. Winograd's system of hints on how to proceed, within his PROGRAMMAR grammar, is, as he himself points out, formally equivalent to an augmented state transition network, and in particular to the ordering of choices at nodes in Woods' system.

There is a significant difference in their metaphysical approaches, or presuppositions about meaning, but it has no influence on the actual operation of their respective systems, and is disguised by the allegiance both give to a procedural view of meaning. The difference is that Woods takes a much more logico-semantic interpretation of that slogan than does Winograd. In particular, for Woods the meaning of an input utterance to his system is the system's procedures for manipulating the truth conditions of the utterance and establishing its truth value. To put the matter crudely, for Woods an assertion has no meaning if his system cannot establish its truth or falsity. Winograd has certainly not committed himself to any such extreme position.

It is interesting to notice that Woods' work is, in virtue of his strong position on truth conditions, probably the only piece of work in the field of AI and natural language to satisfy Hayes' (1974) recent demand that to be "intellectually respectable" a knowledge system must have a natural model theoretic semantics, in Tarski's sense. Since no one has ever given precise truth conditions for any interesting piece of discourse, such as, say, Woods' own papers, one might claim that his theoretical presuppositions necessarily limit his work to the analysis of microworlds (as distinct from everyday language).

There is a low-level problem about the equivalence of Woods' and Winograd's systems, if we consider what we might call the received commonsense view of their work. Consider the following three assertions:

(1) Woods' system is an implementation of a transformational grammar.

(2) Winograd's work has shown the irrelevance of transformational grammar for language analysis--a view widely held by reviewers of his work.

(3) Woods' and Winograd's systems are formally equivalent--a view held by both of them.

There is clearly something of an inconsistent triad among those three widely held beliefs. The trouble probably centers on the exact sense in which Woods' work is formally equivalent to a transformational grammar--not a question that need detain us here, but one worth pointing out in passing.

## Some More General Background Issues

Winograd's work is a central example of the artificial intelligence paradigm of language, using "paradigm" in Kuhn's (1970) sense of a large-scale revision in systematic thinking, where the paradigm revised is the "generative paradigm" of the Chomskyan linguists (Chomsky 1957). From the AI point of view, the generative linguistic work of the last 15 years has three principal defects. First, the generation of sentences, with whatever attached structures, is not in any interesting sense a demonstration of human understanding, nor is the separation of the well-formed from the ill-formed by such methods; understanding requires, at the very least, both the generation of sentences as parts of coherent discourse, and some attempt to interpret, rather than merely reject, what seem to be ill-formed utterances. Neither the transformational grammarians following Chomsky, nor their successors, the generative semanticists (Lakoff 1971), have ever explicitly renounced the generative paradigm.

Second, Chomsky's distinction between performance and competence models, and his advocacy of the latter, have isolated modern generative linguistics from any effective test of the systems of rules it proposes. Whether or not the distinction was intended to have this effect, it has meant that any test situation necessarily involves performance, which is considered to be outside the province of serious linguistic study. Any embodiment of a system of rules in a computer and assessment of its output, would be performance. AI, too, is much concerned with the structure of linguistic processes, *independent of any particular implementation*;* implementation is never excluded, as it is from competence models, but rather encouraged.

Third, as I mentioned before, there was until recently no place in the generative paradigm for inferences from facts and inductive generalizations, even though very simple examples demonstrate the need for it.

---

[*] *Vide*: "Artificial intelligence is the study of intellectual mechanisms apart from applications and apart from how such mechanisms are realized in the human or in animals." (McCarthy, in press.)

This last point, about the shortcomings of conventional linguistics, is not at all new, and in AI is at least as old as Minsky's (1968, p. 22) observation that in "He put the box on the table. Because it wasn't level, it slid off," the last "it" can only be referred correctly to the box, rather than the table, on the basis of some knowledge quite other than that in a conventional and implausible linguistic solution, such as the creation of a class of "level nouns" so that a box would not be considered as being or not being level.

These points would be generally conceded by those who believe there is an AI paradigm of language understanding, but there would be far less agreement over the positive content of the paradigm. The trouble begins with the definition of understanding as applied to a computer. At one extreme are those who say the word can only refer to the performance of a machine--to its ability to, say, sustain some form of dialog long enough and sensibly enough for a human interrogator to be unsure whether what he is conversing with is a machine or not. On the other hand, there are many, almost certainly a majority, who argue that more is required, in that the methods and representations of knowledge by which the performance is achieved must be of the right formal sort, and that mere performance, based on ad hoc methods, does not demonstrate understanding.

This issue is closely related to that of the role of deduction in natural language understanding, simply because deduction is often the structure meant when "right methods" are mentioned. The dispute between those who argue for, or who like Winograd, use deductive methods, and those who advocate other inferential systems closer to commonsense reasoning, is in many ways a pseudo-issue because it is so difficult to define clearly what a nondeductive system is (if by that one means a system that cannot in principle be modeled by a deductive system), since almost any set of formal procedures, including "invalid inferences," can be so displayed. The heart of the matter concerns the most appropriate form of an inference system, rather than how those inferences may be axiomated, and it may well turn out that the most appropriate form for plausible reasoning in order to understand is indeed nondeductive. This same insight has largely defused another heated issue: Whether the appropriate representations should be procedures or declarations. Winograd's work was of the former type, as his definitions of words like "pickup" as procedures for actually picking things up in the blocks world showed. However, simple procedural representations usually have the disadvantage that, if you are going to indicate for every item of knowledge how it is to be used, then, if you may use it on a number of kinds of occasions, you will have to store it that number of times. So, if you want to change it later, you will also have to remember to change it in all the different places you have put it. There is the additional disadvantage of lack of perspicuity: Anyone reading the procedural version of the Winograd grammar rule I gave earlier, will almost certainly find the conventional, declarative version easier to understand.

So then, the fashion for all things procedural has to some extent abated (see Winograd [illegible]). There is general agreement that any system should show, as it were, [illegible] be applied to language, but that is not the same as demand[illegible] be written in a procedural language, like PLANNER. [illegible] last point later.

## Second Generation Systems

To understand what Winograd meant when he contrasted his own with what he called second generation systems, we have to remember, as always in this subject, that the generations are of fashion, not chronology or inheritance of ideas. He described the work of Simmons, Schank, and myself among others in his survey of new approaches, even though the foundations and terminology of those approaches were set out in print in 1966, 1968, and 1967, respectively. What those approaches, and others, have in common is the belief that understanding systems must be able to manipulate very complex linguistic objects, or semantic structures, and that no simplistic approaches to understanding language with computers will work.

In an unpublished, but already very influential recent paper, Minsky (1974) has drawn together strands in the work of Charniak (1972) and the authors above using a terminology of frame:

> A frame is a data-structure for representing a stereotype situation, like a certain kind of living room, or going to a children's birthday party. Attached to each frame are several kinds of information. Some of this is information about how to use the frame. Some is about what one can expect to happen next. Some is about what to do if these expectations are not confirmed.
>
> We can think of a frame as a network of nodes and relations. The top levels of a frame are fixed and represent things that are always true about the supposed situation. The lower levels have many terminals--"slots" that must be filled by specific instances or data. Each terminal can specify conditions its assignments must meet. . . . Simple conditions are specified by markers that might require a terminal assignment to be a person, an object of sufficient value, etc. . . .

The key point about such structures is that they attempt to specify in advance what is going to be said, and how the world encountered is going to be structured. The structures, and the inference rules that apply to them, are also expressions of "partial information" (in McCarthy's phrase) that are not present in first generation systems. As I showed earlier with the "women and soldiers" example such loose inductive information, seeking confirmation from the surrounding context, is required for very simple sentences. In psychological and visual terms, frame approaches envisage an understander as at least as much a looker as a seer.

I shall now describe briefly five approaches that might be called second generation.

### 1. Charniak

The new work that owes most to Minsky's advocacy is Charniak's. He studied what sorts of inferential information (Charniak 1972, 1973, 1974) would be needed to resolve pronoun ambiguities in children's stories, and in that sense to understand them. One of his example stories is: "Jane was invited to Jack's birthday party. She wondered if he would like a kite. A friend told Jane that Jack already had a kite, and that he would make her take it back."

The problem concerns the penultimate word "it," and deciding whether it refers to the first kite mentioned or the second. Charniak's analysis begins by pointing out that a great deal of what is required to understand that story is implicit: knowledge about the giving of presents, knowledge that if one possesses one of a certain sort of thing then one may well not want another, and so on.

Charniak's system does not actually run as a program, but is a theoretical structure of rules called "demons" that correspond roughly to what Minsky later called frames. A demon for this example would be, "If we see that a person might not like a present X, then look for X being returned to the store where it was bought. If we see that happening, or even being suggested, assert that the reason why is that P does not like X."

The important words there are "look for," which suggest that there may well be confirming hints to be found in the story and, if there are, then this tentative, partial inference is correct, and we have a definite and correct answer. This approach, of using partial (not necessarily true) inferences in order to assert a definite answer, is highly characteristic of second generation systems.

The demons are, as with Winograd's work, expressed in a procedural language which, on running, will seek for a succession of interrelated goals.

Here, for example, is a demon concerned with another story, about a child's piggy bank (PB) and a child shaking it, looking for money but hearing no sound. The demon, PB-OUT-OF, is formalized as:

```
(DEMON PB-OUT-OF
        (NOLD PB PERSON M N)
        (?N OUT-OF ?M ?PB)
        (GOAL (? IS ?PB PIGGY-BANK))
        (GOAL (? IS ?M MONEY) $DEDUCE)
        (GOAL (?NOLD SHAKE ?PERSON ?PB)$TRUE)
        (ASSERT (? HAVE ?PERSON ?M))
        (ASSERT (? RESULT ?N ?NOLD)))
```

Again, it is not necessary to explain the notation in detail to see that conditions are being stated for the contents of a piggy bank having been emptied. The pattern being sought by the demon in operation is the third line. If a chain of demons can "reach" one of the possible referents in a story then there is a success registered and the ambiguity of the corresponding pronoun is resolved.

One can see that the information encoded in the system is of a highly specific sort. In the present case it is not about containers as such, and how to get their contents out, but about piggy banks in particular, and everything relies on that particular knowledge having been put in. Not all the knowledge is of this general sort: In a recent paper (Charniak 1974) where the kite story is reconsidered, there are rules of considerable generality and interest. One such is what Charniak calls a R+SSA rule:

> If the story gives information which would make it plausible to infer that PERSON is favorably inclined towards action A, and

PERSON does S, a significant subaction (SSA) of A, then infer that PERSON is doing A.

An important assumption of Charniak's is what I shall call his "decoupling assumption," namely that this work on the knowledge-based determination of reference can be effectively decoupled from any particular syntactic or semantic representation, or of its application, and can be explored in isolation. Thus, the apparently English words in the PB-OUT-OF demon above do not have any particular status other than that of English words that happen to occur in the stories being examined. I shall return to this assumption later.

### 2. Colby

Colby (Colby et al. 1971, Enea and Colby 1973, Colby and Parkison 1974) has developed over a period of several years an online dialog system that simulates a paranoid patient called PARRY. A measure of its success is that it is perhaps the most used AI program in existence; used, that is, by those outside the research project that developed it. Here is a discontinuous sample of its output, again with the system's responses in the upper case:

```
Have you been hospitalized before?

THIS IS THE FIRST TIME.

How long have you been there?

ABOUT TWO WEEKS.

Any headaches?

MY HEALTH IS FINE.

Are you having any difficulty with your memory?

NO.

Then perhaps you can tell me who is the President of the United
States?

NIXON IS PRESIDENT.

What day is it today?

TODAY IS THURSDAY.

What season is it?

LETS GET ON WITH THE INTERVIEW.
```

Colby's aim was to produce a system that was not, in his word, "fragile" in the way that most AI systems are: one that did not break down, or give up, if the input was not exactly what it expected, but one that carried on somehow in all circumstances, in the way that conversing humans do.

Fragility is avoided by having no conventional syntax analyzer, and by not even attempting to take account of all the words in the input. This is a considerable aid, since any parser that begins to parse a more than usually polite request such as "Would you be so kind as to . . ." is going to be in trouble. British English speakers arriving in the United States quickly learn to delete such phrases, since they cause great confusion to human listeners in stores.

The input text is segmented by a heuristic that breaks it at any occurrence of a range of key words. Patterns are then matched with each segment. There are at present about 1,700 patterns on a list (Colby and Parkison, in press) that is stored and matched, not against any syntactic or semantic representations of words (except to deal with contractions and misspellings), but against the input word string direct, and by a process of sequential deletion. So, for example, "What is your main problem" has a root verb "BE" substituted to become

WHAT BE YOU MAIN PROBLEM.

It is then matched successively in the following forms after successive deletions:

BE YOU MAIN PROBLEM

WHAT YOU MAIN PROBLEM

WHAT BE MAIN PROBLEM

WHAT BE YOU PROBLEM

WHAT BE YOU MAIN

and only the penultimate line exists as one of the stored patterns and so is matched. Stored in the same format as the patterns are rules expressing the consequences for the "patient" of detecting aggression and overfriendliness in the interviewer's questions and remarks. The matched patterns found are then tied directly, or via these inference rules, to response patterns which are generated.

Enormous ingenuity has gone into the heuristics of this system, as its popularity testifies. The system has also changed considerably: It is now called PARRY2 and contains the above pattern matching, rather than earlier key word, heuristics. It has partial, or what some would call pragmatic, rules about expectation and intention, and these alone might qualify it as second-generation on some interpretations of the phrase. A generator is also [illegible] installed to avoid the production of only canned responses.

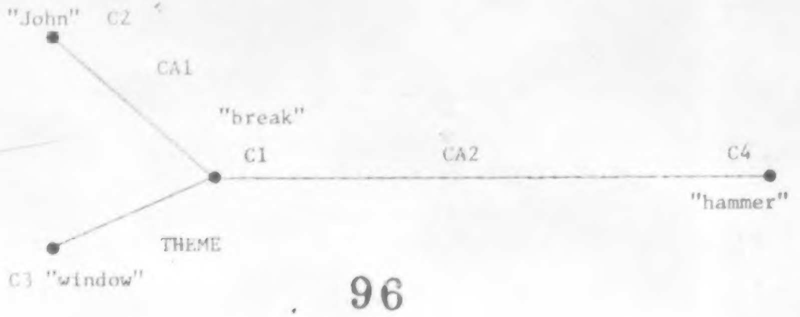
Colby and his associates have put considerable effort [illegible] to find out whether or not psychiatrists can distinguish [illegible] from those of a patient (Colby and Hilf 1973), in what is [illegible] actually to apply Turing's test of machine-person distinction [illegible] are statistical difficulties about interpreting the results [illegible] the result is that the sample questioned cannot distinguish [illegible] whether or not this will influence those who still, on principle, believe that PARRY is not a simulation because it does not understand, remains to be seen. It might be argued that they are in danger of falling into a form of Papert's "human-superhuman fallacy" of attacking machine simulations because they do not perform superhuman tasks, like translate poetry, tasks that some people certainly can do. When such skeptics say that PARRY does not understand, they have in their minds a level of understanding that is certainly high--one could extend their case ironically by pointing out that very few people understand the content of sentences in the depth and detail that an analytic philosopher does, and a very good thing too. There can be no doubt that many people on many occasions do seem to understand in the way that PARRY does.

### 3. Simmons

The remaining three systems differ from the two above in their attempt to provide some representational structure quite different from that of the English input through the use of cases and of complex structures that allow inferences to be drawn from the attribution of case in ways I shall explain. There is also, in the remaining systems, some attempt to construct a primitive, or reduced, vocabulary into which the language represented is squeezed.

Simmons' work is often thought of as a "memory model," though he does in fact pay more attention to word sense ambiguity, and to actual recognition in text than do many other authors. For him the fundamental notion is that of a semantic network, defined essentially by the statement of relational triples of form aRb, where R is the name of a relation and a and b are the names of nodes in the network. Simmons' work with this general formalism goes back to at least 1966 (Simmons et al. 1966) but, in its newer form with case formalism, it has been reported since 1970 (Simmons 1970, Simmons and Bruce 1971, Simmons and Slocum 1972, Simmons 1973, and Hendrix et al. 1973) may reasonably be considered a further implementation of Simmons' methods.

Simmons analyzes the example sentence, "John broke the window with a hammer," into a network of nodes C1, C2, C3, C4 corresponding to the appropriate senses of "John," "break," "window," and "hammer," respectively. The linkages among the nodes are labeled by one of the following "deep case relations": CAUSAL-ACTANT (CA1, CA2), THEME, LOCUS, SOURCE, and GOAL. Case relations are specifications of the way dependent parts of a sentence, or concepts corresponding to parts of a sentence, depend on the main action. So, in this example John is the first causal actant (CA1) of the breaking, the hammer is considered the second causal actant (CA2) of that breaking, and the window is the theme of the breaking. Thus, the heart of the analysis could be represented by a diagram as follows:



or by a set of relational triples:

(C1 CA1 C2)(C1 CA2 C4)(C1 THEME C3)

However, this is not the full representation, and my addition of the word labels to the diagram is misleading, since the nodes are intended to be names of senses of words, related to the actual occurrence of the corresponding word in a text by the relation TOK (for token). In an implementation, a node would have an arbitrary name, such as L97, which would then name a stored sense definition. So, for a sense of "apple" Simmons suggests an associated set of features: NBR-singular (S), SHAPE-spherical, COLOR-red, PRINTIMAGE-apple, THEME-eat, etc. If the name of the node tied to this set of features was indeed L97, then that node might become, say, C5 on being brought into some sentence representation during parsing. Thus, the diagram I gave must be thought of supplemented by other relational ties from the nodes; thus the full sentence about John would be represented by the larger set of triples:

(C1 TOK break)(C1 CA1 C2)(C1 THEME C3)(C1 CA2 C4)
(C2 TOK John)(C2 DET Def)(C2 NBR S)
(C3 TOK Window)(C3 DET Def)(C3 NBR S)
(C4 TOK Hammer)(C4 DET Indef)(C4 NBR S)(C4 PREP With)

Word sense ambiguity is taken account of in that the node for one sense of "hammer" would be different from that corresponding to some other sense of the same word, such as that meaning Edward, Hammer of the Scots, to take a slightly strained alternative for this sentence.*

The network above is also a representation of the following sentences, which can be thought of as surface variants of a single underlying structure:

John broke the window with a hammer

John broke the window

The hammer broke the window

The window broke.

Not all parts of that network will be set up by each of these sentences, of course, but the need for some item to fill an appropriate slot can be inferred; i.e., of the first causal actant (John) in the last two sentences. The sentences above are recognized by means of the "ergative paradigm" of ordered matching patterns, of which the following list is a part:

(CA1 THEME CA2)
(CA1 THEME)
(CA2 THEME)
(THEME)

These sequences will each match, as left-right ordered items, one of the above sentences. It will be clear that Simmons' method of ascribing a node to each word sense is not in any way a primitive system, by which I mean a system of classifiers into which all word senses are mapped.

---

*Simmons' normal example of word sense ambiguity does not apply to the sentence above; he distinguishes "pitcher1," a pouring container, from "pitcher2," in the United States sense of one who throws a ball.

Simmons is, however, considering a system of paraphrase rules that would map from one network to another in a way that he claims is equivalent to a system of primitives. Thus in Simmons (1973) he considers the sentences:

John bought the boat from Mary

Mary sold the boat to John

which would normally be considered approximate paraphrases of each other. He then gives "natural" representations, in his system, as follows in the same order as the sentences:

C1 TOK buy, SOURCE (Mary), GOAL (John), THEME (boat),
C1 TOK sell, SOURCE (Mary), GOAL (John), THEME (boat),

and also the single representation for both sentences, as below, using a primitive action transfer (see description of Schank's work in next section) as follows:

C1 TOK and Args C2, C3
C2 TOK transfer, SOURCE (John), GOAL (Mary), THEME (money)
C3 TOK transfer, SOURCE (Mary), GOAL (John), THEME (boat)

Simmons opts for the first form of representation, given the possibility of a transfer rule going from either of the shallower representations to the other, while Hendrix et al. (1973) adopts the other approach, using a primitive action EXCHANGE instead of transfer.

The implementation under construction is a front-end parser of the Woods augmented transition network type (see Woods 1970), and a generation system going from the semantic networks to surface strings described in detail in Simmons and Slocum (1972). Simmons has also given considerable time (Simmons and Bruce 1971) to the automatic translation of the networks into a corresponding first-order predicate calculus format of the sort developed by Sandewall (1971). This last is particularly valuable because, if generalizable, it shows that any linguistic codings in network form can be translated into some form of the predicate calculus, if that formalism and its associated proof techniques can be shown to be appropriate for certain problems in the area of natural language analysis.

### 4. Schank

Schank's is a rich system of semantic representation, developed over a period of 6 years, with the collaboration of a number of talented students. Its graph system of notation has influenced psychologists like Abelson (1973), among others. Schank's contribution has been the notational system representing the structure of natural language sentences, which has been programed by various collaborators over the years. In its present version, called MARGIE (Schank et al. 1973) it has an analyzer of English due to Riesbeck (1974), a semantic memory component due to Rieger (1974), and a generator of

English due to Goldman (1974). MARGIE produces output in two modes, demonstrating the sort of conceptual inferencing that goes on at the level of the semantic representation:* the PARAPHRASE mode and the INFERENCE mode. Samples of input and output to and from the two modes can be shown thus:

(INFERENCE MODE)

INPUT: John told Mary that Bill wants a book.

OUTPUT: A book about what

Mary knows that Bill wants a book.

Bill wants to come to have a book.

Bill wants someone to cease to have a book.

Bill wants to read a book.

(PARAPHRASE MODE)

INPUT: John killed Mary by choking Mary.

OUTPUT: John strangled Mary.

John choked Mary and she died because she could not breathe.

Mary died because she was unable to inhale some air and she was unable to inhale some air because John grabbed her neck.

The aim of Schank's system has always been to provide a representation of meaning in terms of which these and other tasks, such as machine translation, could be carried out; a representation, moreover, that was to be independent of any particular language, and of syntax, and indeed, of all surface structure whatever.

The formal structure of Shank's graphs is that of dependency grammar (Hays 1964), and the items in the graph are of four types, or conceptual categories. They are symbolized as PP, ACT, PA, and AA, which are acronyms, but which correspond closely (for the purpose of understanding their function) to those of a noun, verb, adjective, and adverb, respectively. (This is a considerable oversimplification, made in order to give a brief and self-contained description. But, in fact, many English nouns are represented as ACTs: chair, pen, honesty, and transportation.) The basic structure is called a conceptualization, and is normally introduced with a straightforward dependency structure such as, for the sentence "The man took a book":

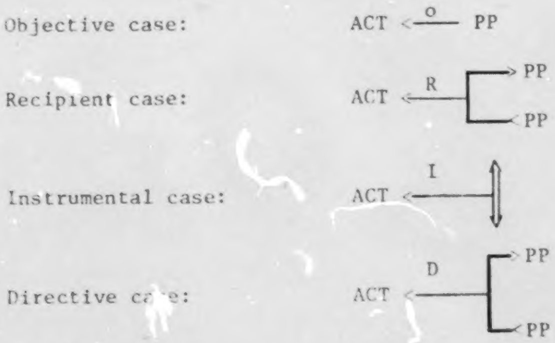
$$\text{Man} \overset{p}{\Longleftrightarrow} \text{take} \overset{o}{\longleftarrow} \text{book}$$

---

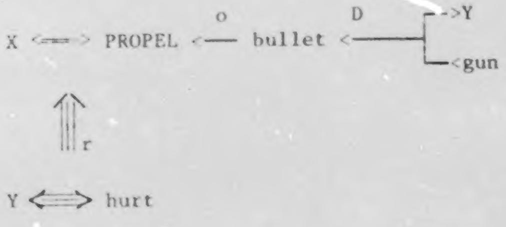
*Schank distinguishes "conceptual" and "semantic" representations in a way that is important to him within his own system. However, I shall use the terms indifferently since, in this brief and superficial description, nothing hangs upon the distinction.

Here "p" indicates past, and <==> is the dependency symbol linking a PP to the ACT ("take") which is the hub of the conceptualization, as with Simmons. The "o" indicates the objective case, marking the dependence of the object PP on the central ACT. There is a carefully constructed syntax of linkages between the conceptual categories, that will be described only in part in what follows.

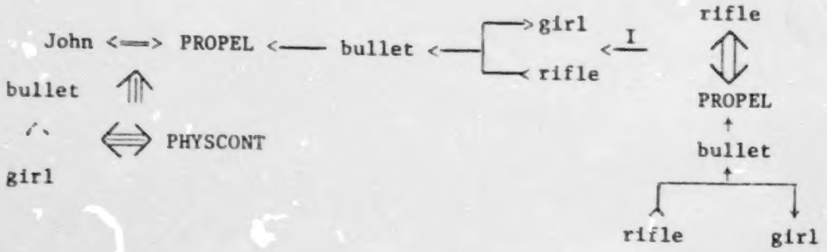
The next stage of the notation involves an extended case notation and a set of primitive ACTs, as well as a number of items such as PHYSCONT which indicate other states, and items of a fairly simplified psychological theory (the dictionary entry for "advise," for example, contains a subgraph telling us that Y "will benefit" as part of the meaning of "X advises Y" (Schank 1973)). There are four cases in the system, and their subgraphs are as follows:



There are at present 14 (Since the publication of Schank (1973) their number has been reduced to 11 (plus DO) by the elimination of SMELL, LISTENTO, LOOKAT, and CONC, and the addition of ATTEND.) basic actions forming the nubs of the graphs, as well as a default action DO. They are: PROPEL, MOVE, INGEST, EXPEL, GRASP, PTRANS, MTRANS, ATRANS, SMELL, SPEAK, LOOKAT, LISTENTO, CONC, and MBUILD. The notions of case and primitive act are related by rules in the development of conceptualizations. So, for example, the primitive act INGEST has as its instrument the act PTRANS. There are also other inferences from any ACT classified as an INGEST action, such as that the thing ingested changes its form; that if the thing ingested is edible the ingester becomes more nourished, etc. (see Schank 1973, pp. 38ff.). This will all become clearer if we consider the transition from a dictionary entry for an action to a filled-in conceptualization. Here is the dictionary entry for the action "shoot":



We can consider this entry as an active framelike object seeking filler items in any context in which it is activated. Thus, in the sentence "John shot the girl with a rifle," the variables will be filled in from context and the case inference will be made from the main act PROPEL, which is that its instrument is MOVE, GRASP, or PROPEL, and so we will arrive at the whole conceputalization:



This case inference must be made, according to Schank, in order to achieve an adequate representation. There is, in the last diagram, a certain redundancy of expression, but as we shall see in the next section this often happens with deeper semantic notations.

More recently, Schank, together with Rieger, has developed a new class of causal inferences that deepen the diagram still further. So, in the analysis of "John's cold improved because I gave him an apple" (in Schank 1974) the extended diagram contains at least four yet lower levels of causal arrowing, including one corresponding to the notion of John constructing the idea (MBUILD) that he wants to eat an apple. So we can see that the underlying explication of meaning here is not only in the sense of linguistic primitives, but in terms of a theory of mental acts as well.

Now there are a number of genuine expositional difficulties here for the commentator faced with a system of this complexity. One aspect is the stages of development of the system itself, which can be seen as a consistent process of producing what was argued for in advance. For example, Schank claimed early on to be constructing a system of semantic structures underlying the surface of natural language, although initially there were no primitives at all, and as late as 1970 (Schank et al.) there was only a single primitive TRANS, and most of the entries in the dictionary consisted of the English words coded, together with subscripts. Since then the primitive system has blossomed and there are now 12 primitives for ACTs including three for the original TRANS itself. Each exposition of the system recounts its preceding phases, from the original primitive-free one, through to the present causal inference form; rather as each human fetus is said to relive in the womb all the evolutionary stages of the human race. The only trouble with this, from an outsider's point of view, is that at each stage the representation has been claimed, in firm tones, to be the correct one, while at the same time Schank admits, in moments of candor (Schank 1973), that there is no end to the conceptual diagraming of a sentence. This difficulty may well reflect genuine problems in language itself, and, in its most acute form concerns a three-way confusion between an attractive notation for displaying the meanings of words, the course of events in the real world, and finally, actual procedures for analysis. It is not always clear whether or not procedures implementing conceptual dependency are intended to recapture all the many phases of expansion of the diagrams.

This raises the important question of the application of a semantic system, that I shall touch on again later. Schank, for example, does mention in passing the questions of word sense ambiguity, and the awful ambiguity of English prepositions, but they are in no way central for him, and he assumes that with the availability of the correct representation, his system when implemented must inevitably solve these traditional and vexing questions. No procedures are hinted at along with the graphs as to how this is to be done. A distinction of importance may be becoming apparent here between Schank's work and Rieger's: In Rieger's thesis (Rieger (1974) the rules of inference appear to create separate and new subgraphs which may stand in an inferential relation to each other so as to produce conclusions about problems of, say, pronoun reference. In Schank's corresponding papers the same inferences are not applied to actual problems (Schank 1974) but only complicate the conceptual graphs further. Closely connected is the question of the survival of the surface structure in the diagrams. Until very recently primitivization applied only to verbs, that of nouns being left to Weber (Weber 1972). Most recently, though, noun words have been disappearing from diagrams and been replaced by categories such as *PHYSOBS*. It is clear that the surface is only slowly disappearing, rather than having been abhorred all along.

In his most recent publication (Schank 1974) there are signs that this trend of infinitely proliferating diagrams is reversing. Schank considers the application of his approach to the representation of text and concludes, correctly in my view, that the representations of parts of the text must be interconnected by causal arrows, and that, in order to preserve lucidity, the conceptual diagrams for individual sentences and their parts must be abbreviated, as by triples such as PEOPLE PTRANS PEOPLE. Here indeed, the surface simply has to survive in the representation unless one is prepared to commit oneself to the extreme view that the ordering of sentences in a text is a purely superficial and arbitrary matter. The sense in which this is a welcome reversal of a trend should be clear, because in the causation inference development, mentioned earlier, all the consequences and effects of a conceptualization had to be drawn within itself. Thus, in the extreme case, each sentence of a text should have been represented by a diagram containing most or all of the text of which it was a part. Thus the representation of a text would have been impossible on such principles.

### 5. Wilks

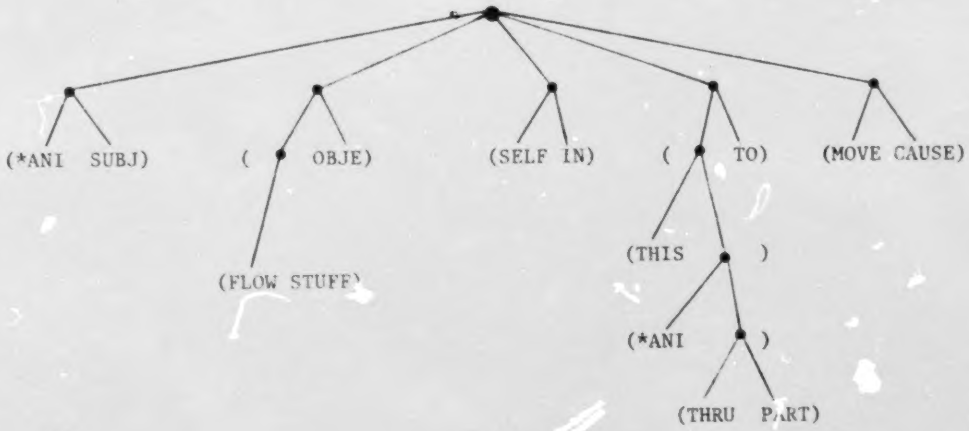
My own system also has a uniform representation, in terms of structures of primitives, for the content of natural language. It is uniform in that information that might conventionally be considered syntactic, semantic, or factual is all represented within a single structure of complex entities called templates, while these are in turn constructed from a budget of 80 primitive semantic entities.

The system runs online as a package of LISP and MLISP programs, taking as input small paragraphs of English, that the user can make up from a vocabulary of about 600 words, and producing a good French translation as output. This environment provides a pretty clear test of language understanding, because French translations for everyday prose are either right or wrong, and can be seen to be so, while at the same time, the major difficulties of understanding programs--word sense ambiguity, case ambiguity, difficult pronoun reference, etc.--can all be represented within a machine translation environment by, for example, choosing the words of the input sentence containing a pronoun reference difficulty so that the possible alternative references have different genders in French. In that way the French output makes quite clear whether or not the program has made the correct inferences in order to understand what it is translating. The program is reasonably robust in actual performance, and will even tolerate a certain amount of bad grammar in the input, since it is not performing a syntax analysis in the conventional sense, but seeking messages representable in the semantic structures employed.

Typical input would be a sentence such as "John lives out of town and drinks his wine out of a bottle. He then throws the bottles out of the window." The program will produce French sentences with different output for each of the three occurrences of "out of," since it realizes that they function quite differently on the three occasions of use, and that the difference must be reflected in the French. A sentence such as "Give the monkeys bananas although they are not ripe because they are very hungry" produces a translation with different equivalents for the two occurrences of "they," because the system correctly realizes, from what I shall describe below as preference considerations, that the most sensible interpretation is one in which the first "they" refers to the bananas and the second to the monkeys, and bananas and monkeys have different genders in French. These two examples are dealt with in the "basic mode" of the system (Wilks 1973). In many cases it cannot resolve pronoun ambiguities by the sort of straightforward preference considerations used in the last example, where, roughly speaking, "ripeness" prefers to be predicated of plantlike things, and hunger of animate things. Even in a sentence as simple as "John drank the wine on the table and it was good," such considerations are inadequate to resolve the ambiguity of "it" between wine and table, since both may be good things. In such cases, of inability to resolve within its basic mode, the program prints COMMON SENSE INFERENCES CALLED on the screen and deepens the representation of the text so as to try and set up chains of inference that will reach, and so prefer, only one of the possible referents. I will return to these processes in a moment, but first I shall give some brief description of the basic representation set up for English.

For each sense of a word in its dictionary the program sees a formula. This is a tree structure of semantic primitives, and is to be interpreted formally using dependency relations. The main element in any formula is the rightmost, called its head, and that is the fundamental category to which the formula belongs. In the formulas for actions, for example, the head will always be one of the primitives PICK, CAUSE, CHANGE, FEEL, HAVE, PLEASE, PAIR, SENSE, USE, WANT, TELL, BE, DO, FORCE, MOVE, WRAP, THINK, FLOW, MAKE, DROP, STRIK, FUNC, or HAPN.

Here is the tree structure for the action of drinking:



Once again, it is not necessary to explain the formalism in any detail, to see that this sense of "drink" is being expressed as a causing to move a liquid object (FLOW STUFF) by an animate agent, into that same agent (containment case indicated by IN, and formula syntax identifies SELF with the agent) and via (direction case) an aperture (THRU PART) of the agent.

Template structures, which actually represent sentences and their parts, are built up as networks of formulas like the one above. Templates always consist of an agent node, an action node, and an object node, and other nodes that may depend on these. So, in building a template for "John drinks wine," the whole of the above tree-formula for "drinks" would be placed at the action node, another tree structure for "John" at the agent node, and so on. The complexity of the system comes from the way in which the formulas, considered as active entities, dictate how other places in the template should be filled.

Thus, the "drink" formula above can be thought of as an entity at a template action node, seeking a liquid object; that is to say a formula with (FLOW STUFF) as its rightmost branch, to put at the object node of the same template. This seeking is preferential, in that formulas not satisfying that requirement will be accepted, but only if nothing satisfactory can be found. The template finally established for a fragment of text is the one in which the most formulas have their preferences satisfied. There is a general principle at work here, that the right interpretation says the least in information-carrying terms. This very simple device is able to do much of the work of a syntax and word sense ambiguity resolving program. For example, if the sentence had been "John drank a whole pitcher," the formula for the "pitcher of liquid" would have been preferred to that for the human, since the subformula (FLOW STUFF) could be appropriately located within it.

A considerable amount of squeezing of this simple canonical form of template is necessary to make it fit the complexity of language: Texts have to be fragmented initially; then, in fragments which are, say, prepositional phrases there is a dummy agent imposed, and the prepositional formula functions as a pseudoaction. There are special, "less preferred" orders to deal with fragments not in agent-action-object order, and so on.

When the local inferences have been made that set up the agent-action-object templates for fragments of input text, the system attempts to tie these templates together so as to provide an overall initial structure for the input. One form is the anaphora tie, of the sort discussed for the monkeys and bananas example above, but the more general form is the case tie. Assignment of these would result in the template for the last clause of "He ran the mile in a paper bag" being tied to the action node of the template for the first clause "He ran the mile," and the tie being labeled CONTainment. These case ties are made with the aid of another class of ordered structures called paraplates, that are attached to the formulas for English prepositions. So, for "outof" there would be at least six ordered paraplates, each of which is a string of functions that seek inside templates for information. In general, paraplates range across two not necessarily contiguous templates. In analyzing "He put the number he thought of in the table," the successfully matching paraplate would pin down the dependence of the template for the last of the three clauses as DIREction, by taking as argument only that one template for the last clause that contained the formula for a numerical table, rather than a kitchen table, in virtue, in this example, of the function in that paraplate seeking a similarity of head (SIGN in this case) between the two object formulas, for "number" and "table."

The structure of mutually connected templates put together thus far constitutes a "semantic block," and if it can be constructed, then as far as the system is concerned, all semantic and referential ambiguity has been resolved and it will begin to generate French by unwrapping the block again. The generation aspects of this work have been described in Herskovits (1973). One aspect of the general notion of preference is that the system should never construct a deeper or more elaborate semantic representation than is necessary for the task in hand and, if the initial block can be constructed and a generation of French done, no "deepening" of the representation will be attempted.

However, many examples cannot be resolved by the methods of this basic mode and, in particular, if a word sense ambiguity, or pronoun reference, is still unresolved, then a unique semantic block of templates cannot be constructed and the extended mode will be entered (Wilks 1973, and in press). In this mode, new templatelike forms are extracted from existing ones and added to the template pool, from which further inferences can be made. In the template derived earlier for "John drinks wine," the system enters the formula for "drinks," and draws inferences corresponding to each case subformula. In this example it will derive templatelike forms equivalent to, in ordinary English, "The wine is in John," "The wine entered John via an aperture," and so on. The extracted templates express information already implicitly present in the text, even though many of them are partial inferences that may not necessarily be true.

Commonsense inference rules are then brought down, which attempt, by a simple strategy, to construct the shortest possible chain of rule-linked template forms from one containing an ambiguous pronoun, say, to one containing one of its possible referents. Such a chain then constitutes a solution to the ambiguity problem, and the preference approach assumes that the shortest chain is always the right one. In the case of "John drank the wine on the table and it was good," the correct chain to "wine" uses the two rules:

I 1. ((*ANI 1)((SELF IN)(MOVE CAUSE))(*REAL 2)) ⟶ (1(*JUDG) 2)

or, in "semi-English,"

[animate-1 cause-to-move-in-self real-object-2] ⟶ [1 *judges 2]

I 2. (1 BE (GOOD KIND)) <⟶ ((*ANI 2) WANT 1)

or, again,

[1 is good] <⟶ [animate-2 wants 1]

These rules are only partial; that is, they correspond only to what we may reasonably look out for in a given situation, not to what must happen. The hypothesis here is that understanding can only take place on the basis of simple rules confirmed by the context of application. In this example the chain constructed may be expressed as [using the above "square bracket notation" to contain not a representation, but simply an indication, in English, of the template contents]:

| | | | |
|---|---|---|---|
| forward inf. | [John drank the wine] | | Template 1 |
| | [John causes-to-move-in-self | wine] | ≑ Template 1 |
| | [John *judges | wine] | by I 1. |
| backward inf. | [John wants | wine] | ≑ line above |
| | [wine is | good] | by I 2. |
| | [?it is | good] | Template 3 |

The assumption here is that no chain using other inference rules could have reached the "table" solution by using fewer than two rules.

The chief drawback of this system is that codings consisting entirely of primitives have a considerable amount of both vagueness and redundancy. For example, no reasonable coding in terms of structured primitives could be expected to distinguish, say, "hammer" and "mallet." That may not matter provided the codings can distinguish importantly different senses of words. Again, a template for the sentence "The shepherd tended his flock" would contain considerable repetition, each node of the template trying, as it were, to tell the whole story by itself. Whether or not such a system can remain stable with a considerable vocabulary of several thousand words has yet to be tested.

---

It will be evident to any reader that the last two systems described, Schank's and my own, share a great deal in common. Even the apparent difference in notation is reduced if one sees the topological similarity that results from considering the head of a formula as functioning rather like a Schank basic action. If one thinks of the dependencies of the case subparts of a formula, not arranged linearly along the bottom of atree. but radiating out from the head in the center, then the two diagrams actually have identical topologies under interpretation. A difference arises in that the "filled-in entity" for Schank is the conceptualization centered on the basic action, though for me it is the network of formulas placed in relation in a template, where there is indeed a basic action (the head of the action formula), but there is also a basic entity in the agent formula and so on. Or, to put it another way, both what-is and what-is-expected are represented in the templates: The agent formula represents the agent, for example, but the left-hand part of the action formula also represents what agent was expected or sought, as in the (*ANI SUBJ) subformula of the "drink" formula.

Although developed in isolation initially, these two systems have also influenced each other in more recent years, probably unconsciously. For example, conceptual dependency now emphasizes the agent-action-object format more than before, and is less "verb-centered" and timeless, while, conversely, my own system now makes much more overt use of rules of partial information than in its earlier versions. Again, both systems have intellectual connections that go back before either generation of AI systems. In my view, both these systems have roots in the better parts of the computational linguistics movement of the fifties: In the case of Schank's system, one may think of the earlier systems of Hays (1964) and Lamb (1966), and the arrow-structured primitive system of Farradene (1966). In the case of my own system there are clear precedents in the Parker-Rhodes (1961) system of classification and the early semantic structures of Richens (1961) and Masterman (1961). In 1961 Masterman was arguing that "what is needed is a discipline that will study semantic message connection in a way analogous to that in which metamathematics now studies mathematical connection, and to that in which mathematical linguistics now studies syntactic connection" (P. 3).

This historical point raises a final one that is, I feel, of passing interest. There seem to be two research styles in this field: One is what might be called the "fully finished style," in which the work exists only in one complete form, and is not issued in early or developed versions. The best example is Winograd's work. The other type, exemplified by all the other authors discussed here, to some extent, is the developing style: work which appears in a number of versions over the years, one hopes with gradual improvements, perhaps in attempts to tackle a wider range of linguistic or other inferential phenomena. There are advantages to both styles, but even in the latter one knows that any proposed structure or system will, in the end, be found wanting in the balances of language. It can only be a question of when one will have to abandon it. The interesting question, and one to which no answer could possibly be given here, is just how far is it worth pushing any given structural approach before starting again from scratch?

## Some Comparisons and Contrasts

In this section I shall compare and contrast the projects described in the body of the paper. A serious problem is that, at this stage of research in artificial intelligence and natural language, the most attractive distinctions dissolve on more detailed scrutiny, largely because of the lack of any precise theoretical statement in most, if not all, the major projects. There are those who think that it therefore follows that this is not the moment for any form of critical comparison in this field, and that no more is needed th[illegible] a positive attitude toward all possible projects. Only those who feel that, on the contrary, any time is as good as any other for the discussion of intellectual differences in the hope of progress, should read on.

One must admit right away that the selection of projects discussed above, like Winograd's distinction between first and second generation systems, on which the selection was to some extent based, cannot be defended by any strict definitions: one that would, in this case, include all the projects described, and exclude all those of Winograd's first generation. One might, for example, want to define second generation systems (in the study of natural language within the AI paradigm) in some very general terms, such as those systems which (1) contain complex semantic structures for the representation of text that are significantly different from the "surface structure" of the input, and (2) contain cognate structures representing conceptual and real world knowledge that is not explicitly present in the input text. Even so general a description of a frames type approach would not cover Charniak or Colby with the first point, nor Simmons with the second, for he has so far eschewed all concern with information not present explicitly in the input text. Moreover, the second point would certainly cover Winograd's own work, as well as other first generation approaches, so it is clear, at the outset of any comparision, that there is not even a simple and unequivocal definition covering all and only the projects to be compared.

## Level of Representation

One important line of current dispute among the second generation approaches concerns the appropriate level of representation for natural language. On the one hand are Colby, and apparently Charniak, who hold that the representation of language can, in effect, be by means of itself. On the other hand, there are Schank and myself, who hold that the appropriate level of computation for inferences about natural language is in some reduced, or primitive, representation. Simmons, as we saw, holds an intermediate position. I wrote "appears" in the case of Charniak because he holds that his structures are independent of any particular level of representation, or rather, that they could be realized at a number of levels of representation, depending on the subject area. However, there is no doubt that the representation in terms of predicates that he offers in his work appears to be in one-to-one correspondence with English words.

The strongest low-level approach is undoubtedly that of Colby, who straightforwardly faces the enormous mapping problems involved if the structures are at the English word level. It is important to realize that this dispute is ultimately one of degree, since no one would claim that every locution recognized by an intelligent analyzer must be mapped into a "deep" representation. To take an extreme case, any system that mapped "Good morning" into a deep semantic representation before deciding that the correct response was also "Good morning" would be making a serious theoretical mistake.

However, the most serious argument for a nonsuperficial representation is not in terms of the avoidance of mapping difficulties, but in terms of theoretical perspicuity of the primitive structures, and this argument is closely tied to the defense of semantic primitives in general, a large subject not to be undertaken here. One of the troubles about semantic primitives is that they are open to bad defenses, which decrease rather than increase their plausibility. For example, some people who use them for linguistic representation have declared them to have some sort of objective existence, and have implied that there is a "right set" of primitives open to empirical discovery. On that view the essentially linguistic character of structures of primitives is lost, for they then might as well be strings of binary numbers, or something equally opaque and nonlinguistic. No great deal of thought is required to see that that simply could not be the case. What is the case is that there is a considerable amount of psychological evidence that people are able to recall the content of what they hear and understand without being able to recall either the actual words or the syntactic structure used. There is ample literature on this subject, from which two sample references would be Wettler (1973) and Johnson-Laird (1974).

These results are, of course, no proof of the existence of semantic primitives, but they are undoubtedly supporting evidence of their plausibility. On a different plane, so is the result from the encoding of the whole Webster's Third International Dictionary at Systems Development Corporation, where it was found that a rank-ordered frequency count of the words used to define other words in that vast dictionary was a list (omitting "the" and "a") which corresponded almost item-for-item to a plausible list of semantic primitives, derived a priori, by those actually concerned to code the structure of word and sentence meanings.

It is important to distinguish the dispute about level from the closely connected topic that I shall call the centrality of the knowledge required by a language understanding system.

## Centrality

What I am calling the centrality of certain kinds of information concerns not its level of representation, but its nonspecificity. Again a contrast can be drawn between the sort of information required by Charniak's system, on the one hand, and that required by Schank's and my own on the other. Charniak's examples suggest that the fundamental form of information is highly specific* to particular situations, like parties and the giving of presents, while the sorts of information central to Schank's and my own systems are general partial assertions about human wants, expectations, and so on, many of which are so general as to be almost vacuous. One might argue that that is why their role in understanding has been ignored for so long.

If I were a reasonably fluent speaker of, say, German, I might well not understand a German conversation about birthday presents unless I had detailed factual information about how Germans organize the giving of presents, which might be considerably different from the way we do it. Conversely, of course, I might understand much of a technical article about a subject in which I was an expert, even though I knew very little of the language in which it was written. These are certainly considerations that tell for Charniak's approach, and it is perhaps a paradox that the sort of natural language understander that would tend to confirm his assumptions would be one concerned with discourse about, say, the details of repairing a motor car, where factual information is what is central, yet, ironically, Charniak has concentrated on something as general as children's stories, with their need of deep assumptions about human desires and behavior.

---

*In Charniak's most recent paper (1974), he gives much more general rules, such as his "rule of significant subaction," mentioned earlier.

In the end this difference may again turn out to be one of emphasis, and of what is most appropriate to different subject areas, though there may be a very general issue lurking somewhere here. It seems to me not a foolish question to ask whether much of what appears to be *about natural language* in AI research is in fact about language at all. Even if it is not, that may in no way detract from its value. Newell (Moore and Newell 1973) has argued that AI work is in fact "theoretical psychology," in which case it could hardly be research *on* natural language. When describing Winograd's work earlier in the paper, *I* raised this question in a weak form by asking whether his definition of "pickup" had anything to do with the natural language use of the word, or whether it was rather a description of how his system picked something up, a quite different matter.

Suppose we generalize this query somewhat, by asking the apparently absurd question of what would be *wrong* with calling, say, Charniak's work an essay on the "Socioeconomic Behavior of American Children Under Stress"? In the case of Charniak's work this is a facetious question, asked only in order to make a point, but with an increasing number of systems in AI being designed not essentially to do research on natural language, but in order to have a natural language "front end" to a system that is essentially intended to predict chemical spectrums, or play snakes and ladders, or whatever, the question becomes a serious one. It seems to me a good time to ask whether we should expect advance in understanding natural language from those tackling the problems head on, or those concerned to build a "front end." It is clearly the case that *any piece of knowledge whatever* could be essential to the understanding of some story. The question is: Does it follow that the specification, organization, and formalization of that knowledge is the study of language, because if it is, then all human inquiry, from physics and history to medicine, is a linguistic enterprise. And, of course, that possibility has actually been entertained within certain strains of modern philosophy.

However, I am not trying here to breathe fresh life into a philosophical distinction between being about language and *not* being about language, but rather introducing a practical distinction (which is also a consideration in favor of opting, as I have, to work on very general and central areas of knowledge), between specific knowledge, and central knowledge *without which a system could not be said to understand the language at all*. For example, I might know nothing of the arrangement of American birthday parties, but could not be accused of not understanding English even though I failed to understand some particular children's story. Yet, if I did not have available some very general partial inference such as the one about people being hurt and falling, or one about people endeavoring to possess things that they want, then it is quite possible that my lack of understanding of quite simple sentences would cause observers to think that I did not understand English. An interesting and difficult question that then arises is whether those who concentrate on central and less central areas of discourse could, in principle, weld their bodies of inferences together in such a way as to create a wider system; whether, to put the matter another way, natural language is a whole that can be built up from parts?

## Phenomenological Level

Another distinction that can be confused with the central-specific one is that of the phenomenological levels of inferences in an understanding system.

I mean nothing daunting by the phrase: Consider the action of eating which is, as a matter of anatomical fact, quite often an act of bringing the bones of my ulna and radius (in my arm) close to that of my lower mandible (my jaw). Yet clearly, any system of commonsense inferences that considered such a truth when reasoning about eating would be making a mistake. One might say that the phenomenological level of the analysis was wrong even though all the inferences it made were true ones. The same would be true of any AI system that made everyday inferences about physical objects by considering their quantum structure.

Schank's analysis of eating contains the information that it is done by moving the hands to the mouth, and it might be argued that even this is going too far from the "meaning" of eating, whatever that may be, toward generally true information about the act, which, if always inferred about all acts of eating, will carry the system unmanageably far.

There is no denying that this sort of information might be useful to have around somewhere; that, in Minsky's terms, the "default" value of the instrument for eating is the hand brought to the mouth, so that, if we have no contrary information, then that is the way to assume that any given act of eating was performed. Nonetheless, there clearly is a danger, and that is all I am drawing attention to here, of taking inferences to a phenomenological level beyond that of commonsense. A clearer case, in my view, would be Schank's analysis (1974) of mental activity in which all actions, such as kicking a ball, say, are preceded by a mental action of conceiving or deciding to kick a ball. This is clearly a level of analysis untrue to commonsense, and which can have only harmful effects in a system intended to mimic commonsense reasoning and understanding.

## Decoupling

Another general issue in dispute concerns what I shall call decoupling, which is whether or not the actual parsing of text or dialog into an understanding system is essential. Charniak and Minsky believe that this initial parsing can be effectively decoupled from the interesting inferential work and simply assumed. But in my view, that is not so, because many of the later inferences would actually have to be done already, in order to have achieved the initial parsing, and so the assumption of decoupling can lead to something like a circularity. For example, in analyzing "He shot her with a colt," we cannot ascribe any structure at all until we can make the inference that guns rather than horses are instruments for shooting.

The inferences required to resolve word sense ambiguities, and those required to resolve pronoun reference problems, are not of different types; often the two problems occur in a single sentence and must be resolved together. But Charniak's decoupling has the effect of completely separating these two closely related linguistic phenomena in what seems to me an unrealistic manner. His system does inferencing to resolve pronoun ambiguities, while sense ambiguity is presumably to be done in the future by some other, ultimately recoupled, system.

Another way of pointing up the difference between the attitudes of second generation systems to decoupling and those of the first generation, is by describing the role of syntax analysis in them. As we saw, syntax was the heart of Winograd's system. Both levels of frame approach discount syntax analysis, though for very different reasons: Charniak does so because it is part of the initial parsing from which his inferential work has been decoupled. Schank and I do so because we believe semantic analysis to be fundamental, and that, in an actual implementation, the results of syntactic analysis can all be achieved by a sufficiently powerful semantic analyzer. This last assumption is confirmed by the limited degree of success that the two semantic analyzers have actually achieved in operation.

## Availability of Surface Structure

An issue close to that of the appropriate level of representation in a system is that of the availability of the surface structure of the language analyzed; or, to put it more crudely, the availability during subsequent analysis of the actual words being analyzed. These are clearly available in Colby, and are indirectly available in Simmons, Winograd, and my own system, but Schank makes a point of the importance of their nonavailability, on the grounds that an ideal representation should be totally independent of the input surface structure and words. There are both theoretical and practical aspects to this claim of Schank's: In the limit, the order of the sentences of a text is part of its surface structure, and presumably it is not intended to abandon this "superficial information." In one of his recent papers (1974) Schank seems to have accepted some limitation on the abandonment of surface structure.

The other, practical, point concerns the form of representation employed: In the (1973) implementation of Schank's system using an analyzer of input text, a memory, and a generator of responses, it was intended that nothing should be transferred from the input program to the output program except a representation coded in the structures of primitives discussed earlier.* The question that arises is, can that structure specify and distinguish word senses adequately without transferring information specifically associated with the input word? Schank clearly believes the answer to this question is yes, but that cannot be considered established by the scale of computations yet described in print.

A suitable environment in which to consider the question is that of translation from one language to another. Suppose we are analyzing a sentence containing the word "nail," meaning a physical object. It is clear that the translation of that word into French should not be the same as the translation for "screw" or "peg." Yet is it plausible that any description of the function of these three entities entirely in terms of semantic primitives, and without any explicit mention of the word name and its connection to its French equivalent, will be sufficient to insure that only the right match is made?

## Application

This point is a generalization of the last two, and concerns the way in which different systems display, in the structures they manipulate, the actual procedures of application of those structures to input text or dialog. This is a matter different from both that of the availability of the surface structure, and of a computer implementation of the system. In the case of Colby's patterns, for example, the form of their application to the input English is clear, even though the matching involved could be achieved by many different implementation algorithms. In the case of my own system, I hold the same to be true of the template structures, even though by the time the input has reached the canonical template form it is considerably different from the input surface structure. The system at the extreme end of any scale of perspicuity of application is Winograd's, where the procedural notation, by its nature, tends to make clear the way in which the structures are applied. At the other end are the systems of Schank and Charniak, where no application is specified, which means that the representations are not only compatible with many implementation algorithms, which does not matter, but are also compatible with many systems of linguistic rules, whose specification is an essential piece of inquiry, and whose subsequent production may cause the basic system to be fundamentally different.

---

*This point is to some extent hypothetical since, as we saw, Schank's conceptualizations still do contain, or appear to contain, many surface items, in particular nouns, adjectives, and adverbs. However, this is a transitional matter and they are in the course of replacement, as noted, by nonsuperficial items.

English prepositions will serve as an example. In Schank's case notation there is no indication of how the case discriminations are actually to be applied to English prepositions in text. So, for example, the preposition "in" can correspond to the containment case, time location, and spatial location, among others. As we saw earlier, the discrimination involved in actual analysis is a matter of specifying very delicate semantic rules ranging over the basic semantic structures employed. Indeed, the structures and case system themselves seem to me to be essentially dependent on the nature and applicability of such rules,* and so this application of the system should have an obvious place in the overall structures. It is not something to be delegated to a mere implementation. If enough of the linguistic intractables** of English analysis were to be delegated out of the representation, AI would be offering no more to the analysis of natural language than the logicians who proffer the predicate calculus as a plausible structure for English.

In some of his more recent writings Winograd has begun to develop a view that is considerably stronger than this application one: In his view the control structure of an understanding program is itself of theoretical significance, for only in that way, he believes, can natural language programs of great size and complexity remain perspicuous.

---

*This is not meant to be just bland assertion. I have written at some length on the relations between application and the theoretical status of linguistic theories in Wilks (1974).

**The differences between Minsky's (1974) notion of "default value" and what I have called "preference" can be pointed up in terms of application. Minsky suggests "gun" as the default value of the instrument of the action of shooting, but I would claim that, in an example like the earlier "He shot her with a colt," we need to be able to see in the structure assigned whether or not what is offered as the apparent instrument is in fact an instrument and whether it is the default or not. In other words, we need sufficient structure of application to see not only that "shooting" prefers an instrument that is gun, but also why it will choose the sense of "colt" that is a gun rather than the one which is a horse.

## Forward Inference

Another outstanding dispute concerns whether one should make massive forward inferences as one goes through a text, keeping all one's expectations intact, as Charniak and Schank hold, or whether, as I hold, one should adopt some "laziness hypothesis" about understanding, and generate deeper inferences only when the system is unable to solve, say, a referential problem by more superficial methods.

Although Schank sometimes writes of a system making "all possible" inferences as it proceeds through a text, this is not in fact the heart of the dispute, since no one would want to defend any strong definition of the term "all possible inferences." Charniak's argument is that, unless certain forward inferences are made during an analysis of, say, a story--forward inferences, that is, that are not problem-driven; not made in response to any particular problem of analysis then known to the system--then, as a matter of empirical fact, the system will not in general be able to solve ambiguity or reference problems that arise later, because it will never in fact be possible to locate (while looking backward at the text, as it were) the points where those forward inferences ought to have been made. This is, in very crude summary, Charniak's case against a purely problem-driven inferencer in a natural language understander.

A difficulty with this argument is the location of an example of text that confirms the point in a noncontentious manner. Charniak has found an excerpt from a book describing the life of apes in which it is indeed hard to locate the reference of a particular pronoun in a given passage. Charniak's case is that it is only possible to do so if one has made certain (nonproblem-occasioned) inferences earlier in the story. But a number of readers find it quite hard to refer that particular pronoun anyway, which suggests that the text was simply badly written.

This is a difficult matter about which to be precise: It would be possible, for example, to agree with Charniak's argument and still construct a purely problem-driven inferencer on the ground that, at the moment, this is the only way one can cope with the vast majority of inferences for understanding, since any system of inferences made in response to no particular problem in the text is too hard to control in practice. Indeed, it is noticeable that the most recent papers of Schank (1974) and Charniak (1974) have been considerably less forward-inference-oriented than earlier ones.*

This dispute is perhaps only one of degree, and about the possibility of defining a degree of forward inference that aids the solution of later semantic problems without going into unnecessary depth.** This might be an area where psychological investigations would be of enormous help to workers in AI.

---

*A particularly interesting withdrawal of a strong forward inference thesis is hidden away on page 283 of Rieger (1974), but has been located by the keen eye of E. Charniak.

**Which may be no more than a psychological restatement of what used to be called (Hayes 1971) (Sandewell 1972) the "frame problem." (No relation, P.E.)

## The Justification of Systems

Finally, one might usefully, though briefly, contrast the different modes of justification implicitly appealed to by the systems described earlier in this paper. These seem to me to reduce to four:

(1) In terms of the power of the inferential system employed. This form of justification has underlain the early predicate calculus-based language programs, and is behind Hayes' (1974) recent demand that any formalism for natural language analysis should admit of a set theoretic semantics, in the Tarskian sense, so as to gain "intellectual respectability," as he puts it. The same general type of justification is appealed to in some degree by systems with PLANNER-type formalisms.

(2) In terms of the provision and formalization, in any terms including English, of the sorts of knowledge required to understand areas of discourse.

(3) In terms of the actual performance of a system, implemented on a computer, at a task agreed to demonstrate understanding.

(4) In terms of the linguistic and/or psychological plausibility of the proffered system of representation.

Oversimplifying considerably, one might say that Charniak's system appeals mostly to (2) and somewhat to (1) and (4); Winograd's to (3) and somewhat to the other three categories; Colby's (as regards its natural language, rather than psychiatric, aspects) appeals almost entirely to (3); Simmons largely to (4), and Schank's and my own to differing mixtures of (2), (3), and (4).

In the end, of course, only (3) counts for empiricists, but there is considerable difficulty in getting all parties to agree to the terms of a test.* A cynic might say that, in the end, all these systems analyze the sentences that they analyze or, to put the same point a little more theoretically, there is a sense in which systems, those described here and those elsewhere, each define a natural language, namely the one to which it applies. The difficult question is the extent to which those many and small natural languages resemble English.

---

*Though an interesting, and potentially revolutionary, distinction seems to have been introduced by a recent reviewer of many of the systems discussed here, between the functioning of a program and a "program in itself": "Only Winograd describes a program that is sufficiently impressive in itself to force us to take his ideas seriously. The techniques of the others have to get by on whatever intuitive appeal they can muster." (Isard 1974)

## CONCLUSION

The last section stressed areas of current disagreement, but there would, if votes were taken, be considerable agreement among AI workers on natural language about where the large problems of the immediate future are: The need for a good memory model has been stressed by Schank (1974), and many would add the need for an extended procedural theory of texts, rather than of individual example sentences, and for a more sophisticated theory of reasons, causes, and motives for use in a theory of understanding. Many might also be persuaded to agree on the need to steer between the Scylla of trivial first generation implementations and the Charybdis of utterly fantastic ones. By the latter, I mean projects that have been seriously discussed but never implemented, for obvious reasons, that would, say, enable a dialog program to discuss whether or not a participant in a given story "felt guilty," and if so, why.

The last disease has sometimes had as a major symptom an extensive use of the word "pragmatics" (though this can also indicate quite benign conditions in other cases), along with the implicit claim that "semantics has been solved, so we should get on with the pragmatics." It still needs repeating that there is no sense whatever in which the semantics of natural language has been solved. It is still the enormous barrier it has always been, even if a few dents in its surface are beginning to appear here and there. Even if we stick to the simplest examples that present no difficulty to the human reader--and one must admit that it has been one of the persistent faults of the AI paradigm of language that it has spent too much time on puzzles examples--there are still great difficulties both systematic and linguistic.

An example of the former would be the development of a dynamic system of understanding texts or stories that had any capacity to recover after having its expectations satisfied and then, subsequently, frustrated. At present no system of the sort described, whether of demons, preference, or whatever, has any such capacity to recover. The situation is quite different from that in a dialog, as in Winograd's system, where, on being given each new piece of information, the system checks it against what it knows, to see if it is being contradicted, and then behaves in an appropriately puzzled way if it is. In frame or expectation systems it is all too easy to construct apparently trick, but basically plausible, examples that satisfy what was being looked for and then overturn it. That possibility is already built into the notion of frame or expectation. A comparison of Phil Hayes' system to my own will serve: Consider "The hunter licked his gun all over, and the stock tasted especially good." What is meant by "stock" is clearly the stock piece of the gun, but any preference system like mine that considers the two senses of "stock," and sees that an edible, soup, sense of "stock" is the preferred object of the action "taste," will inevitably opt for the wrong sense. Any frame or expectation system is prone to the same general kind of counterexample.

In particular cases like this it is easy to suggest what might be done: Here we might suggest a preference attached to the formula for anything that was essentially part of another thing (stock = "part of gun" in this case), so that a local search was made whenever the "part-of" entity was mentioned, and the satisfaction of that search would always be the overriding preference. But that is not the same as a general solution to the problem, which used to be called that of "topic" in the computational semantics of the fifties. There are no solutions to this problem available here and now, though some suggestions have been made by Abelson (1974) and McDermott (1974).

A closely related, but equally intractable, problem is that of how to combine highly specific factual information within a general semantic structuring. Systems like Charniak's are, as we saw, concerned with specific rather than conceptual information, but there are quite simple "semantic specificity" problems that one could not reasonably expect to be tackled even in a system devoted to the handling of facts, as can be seen by contrasting the sentences:

The deer came out of the wood.

The grub came out of the wood.

where we might safely assume that readers would assign quite different senses to "wood" in the two cases simply on the basis of the two different agents. No one, to my knowledge, has suggested any general method for tackling such elementary examples.

But, to finish on the bright side, it is important to stress that there is indeed an AI paradigm of language understanding in existence,* one that embraces first and second generation approaches, and which goes back, I suggested, to a considerable amount of earlier work in computational linguistics. It can be distinguished by a catalog of neglect by conventional linguists that can be summarized under three heads:

(1) Theories of language must have procedural application to the subject matter that could in principle result in computer application and subsequent empirical test.

(2) Theories of language must deal with it in a communicative context, one amenable to empirical assessment. Merely sorting, as generative theories were designed to do, is not enough.

(3) Theories of language must also be, in a clear sense, theories of the formalization and organization of knowledge. If they are not then we can know in advance that they can never tackle the problem of language understanding.

This paper has perhaps overemphasized areas of disagreement and dispute, and undoubtedly many of these disputes will come to be seen as no more than methodological questions, or disputes about matters of emphasis and degree. But some are, I feel sure, questions of substance, and it should be possible to see in the reasonably near future whether one approach on any given question is right and another wrong. It would be nice if this were to be settled by computation rather than by another change of fashion.

---

*One of the very few acknowledgements of this fact, of the possibility of an AI paradigm of language, from a linguist, is found in Fillmore (1974).

# References

In order to compress the reference list the following abbreviations for collections of articles will be used.

| | |
|---|---|
| AIJC3 | Advanced papers of the Third International Joint Conference on Artificial Intelligence. Stanford, Calif., 1973. |
| AISB | Proceedings of the Summer Conference of the Society for Artificial Intelligence and Simulation of Behaviour. University of Sussex, 1974. |
| TEDD | Proceedings of the First International Conference on Machine Translation. National Physical Laboratory, Teddington, 1961 (London: HMSO, 1961). |
| ACL | Proceedings of the Conference of the Association for Computational Linguistics. Amherst, Mass., 1974. |
| CAST | Memorandums from the Istituto per gli studi Semantici e Cognitivi. Castagnola, Switzerland. |
| MITAI | Memorandums from the Artificial Intelligence Laboratory, Massachusetts Insitute of Technology. |
| SUAIM | Memorandums from the Artificial Intelligence Laboratory, Stanford University, Stanford, Calif. |
| SRITN | Technical notes from the Stanford Research Institute, Menlo Park, Calif. |
| MOD | Papers in Computer Models of Thought and Language. R. C. Schank and K. M. Colby, eds. San Francisco: Freeman, 1973. |

Abelson, R. P. The Structure of Belief Systems. In MOD, 1973.

Abelson, R. P. Discourse Headers. To appear in Papers in Memory of J. Carbonell. 1973.

Balzer, R., et al. Domain-Independent Automatic Programming. USC/Information Sciences Institute, Marina del Rey, California, 1974.

Bobrow, D. G. Natural Language Input for a Computer Problem-Solving System. In M. Minsky (ed.) Semantic Information Processing. Cambridge: MIT Press, 1968.

Brown, G. The Believer System. In ACL, 1974.

Bruce, B. A model for temporal references and its application in a question answering program. In Artificial Intelligence 3, 1972.

Bruce, B., and Schmidt, C. F. Episode Understanding and Belief Guided Parsing. In ACL, 1974.

Charniak, E. Towards a Model of Children's Story Comprehension. MITAI-TR 266, 1972.

Charniak, E. Jack and Janet in Search of a Theory of Knowledge. In AIJC3, 1973.

Charniak, E. "He will make you take it back": A Study in the Pragmatics of Language. CAST No. 5, 1974.

Chomsky, N. Syntactic Structures. The Hague: Mouton, 1957.

Colby, K. M. and Parkison, R. C. Pattern-matching rules for the recognition of natural language dialogue expressions. American Journal of Computational Linguistics, 1 (1), 1974.

Colby, K. M. and Hilf, F. D. Multidimensional Analysis in Evaluating a Simulation of Paranoid Thought. In SUAIM-194, 1974.

Colby, K. M., et al. Artificial paranoia. Artificial Intelligence 2, 1971.

Coles, L. S. Techniques for Information Retrieval Using an Inferential Question-Answering System with Natural Language Input. SRITN, 1972.

Davidson, D. Actions, reasons and causal laws. Journal of Philosophy, 1963.

Davies, D. J. M. and Isard, S. D. Utterances as Programs. Presented at the Seventh International Machine Intelligence Workshop, Edinburgh, Scotland, June 1972.

Enea, H. and Colby, K. M. Idiolectic Language-Analysis for Understanding Doctor-Patient Dialogues. AIJC3, 1973.

Farradene, B. Research on Information Retrieval by Relational Indexing. Department of Information Sciences, City University, London, 1966.

Fillmore, C. Summer School Lectures. University of Pisa, Italy, 1974.

Goldman, N. M. Computer Generation of Natural Language from a Deep Conceptual Base. CAST No. 2, 1974.

Halliday, M. A. K. In Foundations of Language, 1970.

Hayes, P. J. The Frame Problem and Related Problems in Artificial Intelligence. In SUAIM-53, 1971.

Hayes, P. J. Some Problems and Non-Problems in Representation Theory. In AISB, 1974.

Hays, D. Dependency theory: A formalism and some observations. Language, 40, 1964.

Heidorn, G. E. English as a Very High Level Language for Simulation Programming. In Proc. Symposium on Very High Level Languages, Santa Monica, California, 1974.

Hendrix, G. G., et al. Language Processing via Canonical Verbs and Semantic Models. In AIJC3, 1973.

Herskovits, A. The Generation of French from a Semantic Representation. SUAIM-212, 1973.

Hewitt, C. In Proc. Second International Joint Conference on Artificial Intelligence, Bedford, Massachusetts: Mitre Corp., 1969.

Hockett, C. F. Grammar for the Hearer. In Vol. XII of the Symposia in Applied Mathematics, American Mathematical Society, 1962.

Isard, S. Review of MOD. Science, 186, 1974.

Johnson-Laird, P. Memory for words. Nature, 251, 1974.

Joshi, A. K. and Weischedel, R. M. Some Frills for the Modal Tic-Tac-Toe of Davies and Isard: Semantics of Predicate Complement Construction. In AIJC3, 1973.

Kuhn, T. S. The Structure of Scientific Revolutions, 2d ed. Chicago: University of Chicago Press, 1970.

Lakoff, G. In Steinberg and Jakobovits (eds.), Semantics. (CUP, Cambridge, 1971).

Lamb, S. In Outline of Stratificational Grammar, Washington, D. C.: Georgetown University Press, 1966.

Lighthill, J. In Artificial Intelligence: A Paper Symposium, Science Research Council, 1973.

McCarthy, J. and Hayes, P. J. In Meltzer and Michie (eds.), Machine Intelligence 4, Edinburgh: Edinburgh University Press, 1969.

McCarthy, J. Review of J. Lighthill, Artificial Intelligence: A Paper Symposium. To appear in Artificial Intelligence.

McDermott, D. V. Assimilation of New Information by a Natural Language Understanding System. MITAI, Tr-291, 1974.

Masterman, M. Semantic Message Detection for Machine Translation. In TEDD, 1961.

Minsky, M. L. Editor's Introduction to Semantic Information Processing. Cambridge: MIT Press, 1968.

Minsky, M. Frame systems. Unpublished MSS. From Massachusetts Institute of Technology, 1974.

Moore, J. and Newell, A. How can Merlin Understand? In L. Gregg (ed.), Knowledge and Condition. Potomac, Maryland: L. Erlbaum Association, 1973.

Parker-Rhodes, A. F. A New Model of Syntactic Description. In TEDD, 1961.

Richens, R. H. Tigris and Euphrates. In TEDD, 1961.

Rieger, C. J. Conceptual Memory. Unpublished doctoral dissertation, Stanford University, 1974.

Riesbeck, C. K. Computational Understanding: Analysis of Sentences and Context. CAST No. 4, 1974.

Rulifson, J. A., et al. QA4: A Procedural Calculus for Inductive Reasoning. SRITN, 1973.

Sandewall, E. Representing Natural Language Information in Predicate Calculus. In Meltzer and Michie (eds.), Machine Intelligence 6. Edinburgh: Edinburgh University Press, 1971.

Sandewall, E. An Approach to the Frame Problem, and its Implementation. In Meltzer and Michie (eds.), Machine Intelligence 7. Edinburgh: Edinburgh University Press, 1972.

Sandewall, E. PCF-2, A First-Order Calculus for Expressing Conceptual Information. Department of Computer Science, Uppsala University, Uppsala, Sweden, 1972.

Schank, R. C. Semantic Categories. Tracor Doc. No. 68-551-U, 1968.

Schank, R. C. The Fourteen Primitive Actions and Their Inferences. SUAIM-183, 1973.

Schank, R. C. Identification of Conceptualizations Underlying Natural Language. In MOD, 1973.

Schank, R. C. Causality and Reasoning. CAST No. 1, 1974.

Schank, R. C. Understanding Paragraphs. CAST No. 6, 1974.

Schank, R. C., et al. SPINOZA II: Conceptual Case-Based Natural Language Analysis. SUAIM-109, 1970.

Schank, R. C., et al. Primitive Concepts Underlying Verbs of Thought. SUAIM-162, 1972.

Schank, R. C. and Rieger, C. J. Inference and the Computer Understanding of Natural Language. SUAIM-197, 1973.

Schank, R. C., et al. MARGIE: Memory, Analysis, Response Generation and Inference on English. In AIJC3, 1973.

Simmons, R. F. Natural Language Question-Answering Systems. Comm. ACM, 13, 1969.

Simmons, R. F. Some Semantic Structures for Representing English Meanings. Tech. Report No. NL-1. Austin, University of Texas, 1970.

Simmons, R. F. Semantic Networks: Their Computation and Use for Understanding English Sentences. In MOD, 1973.

Simmons, R. F., et al. An Approach Toward Answering English Questions from Text. In *IFIPS Conf. Proc., FJCC*, 29, 1966.

Simmons, R. F. and Bruce, B. C. Some Relations between Predicate Calculus and Semantic Net Representations of Discourse. In *Proc. Second International Conference on Artificial Intelligence*, 1971.

Simmons, R. F. and Slocum, J. Generating English Discourse from Semantic Networks. In *Comm. ACM*, 15, 1972.

Sussman, G. J. The Virtuous Nature of Bugs. In AISB, 1974.

Weber, S. Semantic Categories of Nominals for Conceptual Dependency Analysis of Natural Language, SUAIM-172, 1972.

Wettler, M. Zur Speicherung syntaktischer Merkmale in Langzeitgedächtnis. *Bericht über den 28. Konress der Deutschen Gesellschaft f. Psychologie*, 1973.

Wilks, Y. A. Semantic Consistency in Text--an Experiment and some Suggestions, Systems Development Corp., Santa Monica. California, SP-2758, 1967

Wilks, Y. A. *Grammar, Meaning and the Machine Analysis of Language*, Boston and London: Routledge, Kegan Paul Ltd., 1972.

Wilks, Y. A. Preference Semantics, in SUAIM-206, and to appear in Keenan (ed.), *Formal Semantics of Natural Language*, Cambridge: Cambridge University Press, 1975.

Wilks, Y. A. An Artificial Intelligence Approach to Machine Translation, in MOD, 1973.

Wilks, Y. A. One Small Head--Models and Theories in Linguistics, in *Foundations of Language*, 11, 1974.

Wilks, Y. A. A Preferential, Pattern Seeking, Semantics for Natural Language Inference, to appear in *Artificial Intelligence*.

Winograd, T. Procedures as a Representation for Data in a Computer Program for Understanding Natural Language, MIT Ph.D. Thesis, 1971.

Winograd, T. *Understanding Natural Language*. New York: Academic Press, Inc., 1972.

Winograd, T. *A Survey Lecture on Natural Language*, Third International Joint Conference on Artificial Intelligence, Stanford, California, 1973.

Winograd, T. Frame Representation and the Declarative/Procedural Controversy, to appear in Essays in Memory of Jaime Carbonell, 1974.

Winograd, T. Five Lectures on Artificial Intelligence, SUAIM-246, 1974.

Wittgenstein, L. Philosophical Investigations, Blackwell, Oxford, 1953.

Woods, W. Transition Network Grammars for Natural Language Analysis, Comm. ACM, 13, 1970.

Woods, W. Meaning and Machines in Zampolli (ed.), Proc. of the International Conference on Computational Linguistics, Pisa, 1973.

Woods, W., et al. The Lunar Sciences Natural Language Information System, BBN Report 2378, 1972.

☆ U.S. GOVERNMENT PRINTING OFFICE: 1976— 623-244/490